

# Open Research Online

---

The Open University's repository of research publications and other research outputs

## Problems of evaluation with special reference to the evaluation of Open University counselling

### Thesis

#### How to cite:

Thomas, Alan (1977). Problems of evaluation with special reference to the evaluation of Open University counselling. PhD thesis The Open University.

For guidance on citations see [FAQs](#).

© 1977 The Author



<https://creativecommons.org/licenses/by-nc-nd/4.0/>

Version: Version of Record

Link(s) to article on publisher's website:

<http://dx.doi.org/doi:10.21954/ou.ro.0000fca4>

---

Copyright and Moral Rights for the articles on this site are retained by the individual authors and/or other copyright owners. For more information on Open Research Online's data [policy](#) on reuse of materials please consult the policies page.

---

[oro.open.ac.uk](http://oro.open.ac.uk)

D37921/81

UNRESTRICTED

THE OPEN UNIVERSITY OF GREAT BRITAIN

PROBLEMS OF EVALUATION WITH SPECIAL REFERENCE  
TO THE EVALUATION OF OPEN UNIVERSITY COUNSELLING

A THESIS SUBMITTED  
IN CANDIDACY FOR THE DEGREE OF  
DOCTOR OF PHILOSOPHY

BY

ALAN BERKELEY THOMAS, B.A.,

EDUCATIONAL TECHNOLOGY

FEBRUARY 1977

Date of Submission: 3.3.77  
Date of award: 28.10.77

OPEN UNIVERSITY LIBRARY



ProQuest Number: 27777454

All rights reserved

INFORMATION TO ALL USERS

The quality of this reproduction is dependent on the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if material had to be removed, a note will indicate the deletion.



ProQuest 27777454

Published by ProQuest LLC (2020). Copyright of the Dissertation is held by the Author.

All Rights Reserved.

This work is protected against unauthorized copying under Title 17, United States Code  
Microform Edition © ProQuest LLC.

ProQuest LLC  
789 East Eisenhower Parkway  
P.O. Box 1346  
Ann Arbor, MI 48106 - 1346

# ABSTRACT

The thesis is concerned with problems of evaluation, both in the general sense of how anything may be evaluated and more specifically with the topic of the evaluation of the Open University counselling service.

Although evaluation has long been a topic of concern to philosophers, the specialism of educational evaluation has emerged relatively recently. In particular, a body of 'theory' and practice has developed in the United States, oriented initially to the evaluation of curricula and increasingly towards the servicing of educational decisions of all kinds. A number of schemes for the conduct of evaluation have been proposed, yet there is widespread dissatisfaction with the results of evaluative research.

Given the problematic nature of evaluation in education, and that an examination of the concept of evaluation might throw some light on the difficulties, it is necessary to consider some of the philosophical positions on the subject. Following this a model of evaluation is outlined, and the importance of standards for the process of evaluation is emphasised. It is argued that calls for formal evaluations typically arise under conditions of dispute among decision-makers, and some possible ways of resolving such disputes are considered. The need to establish the decision-makers' degree of incorrigibility with regard to research findings is also proposed.

An evaluation project on the Open University counselling service faced a number of difficulties, and the progress of this work is reported. The account provides a relatively rare glimpse of the social and organisational context of evaluative research in a highly unusual institution. The problems are considered in the light of the analysis of the evaluation process, and in terms of the specific circumstances surrounding the project.

Finally, some suggestions are made as to how more fruitful evaluations might be achieved in the future.

## CONTENTS

	PAGE
ACKNOWLEDGEMENTS .....	8
LIST OF TABLES .....	10
LIST OF FIGURES .....	11
LIST OF SUPPLEMENTARY MATERIAL .....	12
CHAPTER ONE. INTRODUCTION .....	13
CHAPTER TWO. THE RISE OF EVALUATION IN EDUCATION .....	33
Educational Evaluation in the United States .....	34
Evaluation Under the Elementary and Secondary Education Act .....	40
Evaluation and the Accountability Movement .....	43
Summary of Developments in the United States .....	45
Educational Evaluation in Britain .....	46
Summary and Comments .....	50
CHAPTER THREE. PRACTITIONERS' VIEWS OF EVALUATION .....	61
N.E. Gronlund: Measurement and Evaluation in Teaching, 1968.....	62
M. Parlett and D. Hamilton, Evaluation as Illumination, 1972 .....	63
M. Provus, Discrepancy Evaluation, 1971 .....	65

	PAGE
M. Scriven, The Methodology of Evaluation, 1967 .....	67
D.L. Stufflebeam, <u>et al</u> , Educational Evaluation and Decision-Making, 1971 .....	69
E.A. Suchman, Evaluative Research, 1967 .....	71
C.R. Wright, Evaluation Research .....	73
R.E. Stake, The Countenance of Educational Evaluation, 1967 .....	75
Comments on Practitioners' Views .....	79
CHAPTER FOUR. THE NATURE OF EVALUATION .....	87
A Philosophical Consideration of Evaluation .....	90
Some Positions in Philosophical Ethics .....	95
Preliminary Considerations of the Nature of Evaluation..	101
The Need for Evaluation .....	112
Summary of Philosophical Considerations.....	114
CHAPTER FIVE. A MODEL OF EVALUATION .....	122
Implicit and Explicit Evaluation .....	122
An Evaluation Model .....	125
Subject Variation .....	132
Constitutive and Associative Variables .....	133
The Set of Significant Descriptions .....	136
Evaluative Terms .....	145
The Evaluative Function .....	150
An Overview of the Model .....	153
Evaluating Courses of Action .....	161
Sources of Standards .....	166
Justification and Regress .....	168
Some Features of Values and Preferences .....	173
Points of View .....	176

CONTENTS (cont'd)

	PAGE
Levels of 'Sophistication' in Evaluation .....	177
Standards for Standards .....	186
The Nature of Evaluative Research in Education .....	191
CHAPTER SIX. EVALUATIVE DISPUTES .....	204
Resolving Evaluative Disputes .....	223
The Contentious Context of Evaluative Research .....	229
The Adversary Approach .....	243
CHAPTER SEVEN. THE EVALUATION OF OPEN UNIVERSITY	
COUNSELLING: I: THE BACKGROUND .....	249
The Background of Opinion .....	254
Changes in the System .....	269
Comments .....	278
CHAPTER EIGHT. THE EVALUATION OF OPEN UNIVERSITY	
COUNSELLING: II: TUITION AND COUNSELLING RESEARCH	
1970-73 .....	286
1970: The Beginnings of the Project .....	286
The Project in 1971 .....	293
The Project in 1972 .....	303
The Project in 1973 .....	308
CHAPTER NINE. PROBLEMS OF EVALUATING OPEN UNIVERSITY	
COUNSELLING .....	325
Difficulties in Evaluating Counselling .....	332
Difficulties Concerning Standards .....	333
Difficulties in Obtaining Information .....	339
Additional Difficulties of the Project .....	354
Summary of Problems .....	359
Some Comments on Evaluation and Decision-Making .....	361

CONTENTS (Cont'd)

	PAGE
CHAPTER TEN. SUMMARY AND CONCLUSIONS .....	387
Summary .....	387
Some Suggestions for the Future .....	393
APPENDIX I. TUITION AND COUNSELLING RESEARCH PROJECTS INVOLVING THE AUTHOR .....	401
APPENDIX II. CHRONOLOGICAL SUMMARY OF DEVELOPMENTS IN THE TUITION AND COUNSELLING SYSTEM .....	405
APPENDIX III. TUITION AND COUNSELLING ARRANGEMENTS .....	407
APPENDIX IV. OBJECTIVES OF TUITION AND COUNSELLING RESEARCH, 1971 .....	411
APPENDIX V. RESEARCH INTO COUNSELLORS AND COUNSELLING IN 1973 .....	417
APPENDIX VI. 'TUITION AND COUNSELLING RESEARCH: WHAT SHOULD WE BE DOING?' BY R.A. NORTHEGE .....	425
BIBLIOGRAPHY .....	435

### ACKNOWLEDGEMENTS

There are a number of people I would like to thank for their help and advice in the preparation of this work.

First and foremost my thanks to my supervisor, Prof. B.N. Lewis, Deputy Director of the Open University Institute of Educational Technology. He has given me numerous helpful insights into my topic, and his constructive criticisms and suggestions have enabled me to amend the thesis at a number of points. Without his support, encouragement and advice, this thesis would never have been started, and probably not have been completed.

I also wish to thank Prof. D.G. Hawkridge, Director of the Open University Institute of Educational Technology. I worked as a Research Assistant under his direction for most of my time at the Open University, and was able to benefit from his experience and his views on evaluation matters. He also read a draft of the thesis and I am grateful for his encouraging remarks and his tolerance.

Prof. Gordon Pask, of Brunel University, and his assistant, Dionysius Kallikourdis also read the manuscript and made valuable comments.

My thanks also to my former colleague on the Tuition and Counselling Research Project at the Open University, Andrew Northedge. Many hours were spent discussing research problems together and I



learnt much as a result. I am particularly grateful for his permission to reproduce in full his paper 'Tuition and Counselling Research: What Should We Be Doing?'.

I also wish to thank David Harris, a former colleague at the Institute of Educational Technology; Alan Crispin, an Open University Staff Tutor; and all the other members of the Open University who, directly or indirectly, have given me help or have provided me with 'raw materials' in the form of their own behaviour, reports, or contributions as subjects of my own researches.

No thesis can reach completion without the diligent efforts of a typist, and so my thanks to Stella Leskiw and Barbara Dinwoodie for the excellent work they have done.

Finally, I must acknowledge the great debt I owe to my wife for her support. It has been a long and sometimes painful journey, and it is she who has had to put up with me. For this I offer my heartfelt thanks.

A.B.T.

LIST OF TABLES

	PAGE
1. Summary of Practitioners' Views .....	78
2. Envisaged Functions of Counsellors in 1969 .....	262
3. Explanations of Difficulties Associated with the Study Centre Visits Project .....	300
4. Proposed Projects for Tuition and Counselling Research for 1972 .....	305
5. Proposed Projects for Tuition and Counselling Research for 1973 .....	309
6. Criterion Variables for Counsellors Identified by the Critical Incidents Method .....	331
7. Report and Record Forms Used in the Counselling System .....	341
8. Frequency of Attendance at Study Centres Claimed by Early Drop-outs and Non-continuers, 1971 .....	346
9. Reasons for Low or Non-attendance at Study Centres by Early Drop-outs and Non-continuers, 1971 .....	346
10. Satisfaction with the Tuition and Counselling Arrangements Claimed by Early Drop-outs, 1971 .....	347
11. Sources Asked for Help with the Course Materials by Early Drop-outs and Whether Found Helpful, 1971 ...	347
12. Persons with Whom Early Drop-outs and Non-continuers Discussed Withdrawal, 1971 .....	348
13. Satisfaction with the Counselling Service Claimed by Non-continuers, 1971 .....	348
14. Satisfaction with the Tuition and Counselling Arrangements Claimed by Students Completing Courses in 1972 .....	349
15. Main Reasons for Dissatisfaction with Counselling Given by Students Completing Courses in 1972 .....	349

LIST OF FIGURES

	PAGE
1. Some Features of the Open University and Its Organisation.....	21
2. Branches of the Philosophy of Ethics .....	93
3. The Evaluation Process with an Established Set of Evaluative Standards.....	131
4. Pascal's Triangle and the Number of Ways of Ordering Given Numbers of Descriptions .....	139
5. Pascal's Triangle in Grid Form .....	151
6. Hypothetical Set of Standards for Counsellors .....	162
7. Hypothetical Justification for a Standard for Counsellors .....	172
8. Levels of 'Sophistication' of Evaluative Schemes.....	181
9. Summary of Tasks for Explicit Evaluations .....	190
10. Configurations of Agreement and Dispute Within Standards .....	215
11. The 1971 Tuition and Counselling System .....	264
12. The Tuition and Counselling System, 1972-75 .....	272
13. Some Hypothetical Standards for the Counselling Service .....	353

LIST OF SUPPLEMENTARY MATERIAL

1. Tuition and Counselling Research Project: Summary Report on Phase 1a (1971).
2. Tuition and Counselling Research Project: Report on Phase 1b (1971).
3. Successful Counselling Practices: Tuition and Counselling Research Project: Phase 2a (1971).
4. Reports of Visits to Study Centres: Tuition and Counselling Research Project: Phase 2a (1971).
5. Analysis of Form C4: Counsellors' Report on Counselling and the Study Centre: Preliminary Report (1971).
6. Final Report on the Analysis of Counsellors' Form C4 (1972).
7. Tuition and Counselling Research Phase 1b, 1972: An Examination of Dropout, Wastage and Conservation of Part-Time Tutors and Counsellors 1971-72 (1972).
8. Tuition and Counselling Research Phase 1a, 1972: Census of 1972 Part-time Staff (1972).
9. Report on the 1973 Counsellor Characteristics Study (1974).

## CHAPTER ONE

### INTRODUCTION

The topic of this thesis is 'Problems of Evaluation with Special Reference to the Evaluation of Open University Counselling'. The contents can be divided roughly into two parts. The first part is concerned mainly with the clarification of the nature of evaluation, and leads to the specification of a model of the evaluative process. The second part consists of an account of an evaluation research project into the counselling services provided by the Open University of Great Britain, undertaken between 1971 and 1974. The parts are linked via an interpretation of the problems and progress of the research project in terms of the elements of the evaluation model. The intention, therefore, is to show:

- a) What needs to be done in order to evaluate something (more precisely, to evaluate something explicitly)
- b) What was done when a particular attempt was made to evaluate the Open University's counselling services
- c) To illustrate some of the difficulties attendant upon evaluation by considering this specific evaluative problem in relation to the demands of the evaluation model

In fulfilling these intentions, it is hoped to provide a better understanding of the problems which surround evaluation.

The thesis might be thought to be rather unusual in so far as it does not report on the conduct and findings of a single piece of research. However, the topic dealt with emerged from the research experiences the writer gained as a Research Assistant in the Open University's Institute of Educational Technology over a period of three years. During this time the writer conducted more than ten separate empirical investigations associated with the University's tuition and counselling services, ranging from the statistical analysis of rates of 'turnover' among part-time staff, to a 'critical incidents' investigation of counselling using tape-recorded interviews and questionnaires. A twenty-five page paper on the latter project was published in Teaching At A Distance in November 1974;<sup>1</sup> references to other relevant research reports are given in subsequent chapters, and a list of the studies with which the author was associated is given in Appendix I. Copies of selected reports and papers produced by the author have been submitted under separate covers to accompany the thesis.

It was this research work which gave the initial stimulus for the present inquiry. As a result of trying to do evaluative research, the writer was encouraged to ask a number of conceptual questions about evaluation, and to consider the nature of the process of evaluation. Although we do not focus on a single study here, the ideas presented could probably not have been produced without this experience. This is not to say that no empirical material is included. Rather

we have reported on the progress of a whole research project consisting of a number of different studies.

It is customary, according to Parsons,<sup>2</sup> for the writers of theses to refer to the subject of their work as 'this strangely neglected topic'. The subject of evaluation is, however, one which has received attention from thinkers for over two thousand years. Philosophers, for instance, have been pondering the nature of morality, the Good, and so forth, since at least the time of Plato, and Ethics has long been one of the major branches of philosophical inquiry. But despite this long history of investigation, evaluation remains a puzzling problem from the point of view of the discipline of philosophy. On the one hand, there are those who have dismissed evaluative phenomena as irrational and meaningless; whilst on the other, reside a variety of ethical theories - objectivist, subjectivist, emotivist, naturalistic, etc. The existence of these differing positions demonstrates that there is no consensus among philosophers about the nature of evaluation.

This thesis is not, however, intended as a treatise in philosophy. Rather the insights of philosophers of ethics have been used to inform a newer and more pragmatic field, that of educational evaluation. It is not easy to give a precise definition of educational evaluation, for, as we shall see, a number of different approaches to the topic are offered by practitioners. Broadly speaking, it can be regarded as an activity which harnesses social science research techniques to the service of educational decision-making and the control of educational organisations. It can

therefore be considered as one application of evaluative research in general, which finds other applications in fields such as public health, business and industry.

Educational evaluation, in its institutionalised and formalised form, is certainly a recent enterprise. A fuller resumé of the development of educational evaluation is given in the next chapter, but it is sufficient to point out here that it is probably only in the last fifteen years or so that it has become Big Business. The field is more 'developed' in the United States, but even in Britain the practice and profession of educational evaluation is becoming more common. Sometimes the focus of evaluation is seen as educational curricula,<sup>3</sup> sometimes as virtually anything connected with the running of an ongoing educational institution.<sup>4</sup> In any event, a literature has developed on the subject which offers numerous conceptual distinctions and which proposed a number of models, theories, procedural specifications and recipes for doing educational evaluation.

It is perhaps typical that a newly emerging field should generate a diverse and heterogeneous literature. Philosophers of science would no doubt attribute this to the field's immaturity. Whatever the cause, the overall impression that is obtained from an attempt at a synthetic appreciation of the educational evaluation literature is one of confusion. As Parlett and Hamilton put it:



"Confusion is engendered as rival proposals, models and terminologies are voiced and then rapidly countered. As a developing field of study, evaluation proceeds in the absence of coherent and agreed frames of reference."<sup>5</sup>

At the same time, there appears to be a general dissatisfaction with the way in which educational evaluation is conducted and with the results it achieves. Stufflebeam,<sup>6</sup> for instance, devotes a whole chapter to an 'Evaluation of Evaluation' in which he concludes that evaluation is in a pathological state. Clients of evaluative research are anxious, sceptical, and eager to avoid involvement in evaluative studies, whilst practitioners give bad advice, obtain insignificant results, are poorly trained, and lack relevant theories, guidelines and techniques. And according to Guttentag, "Both practitioners and consumers of evaluation are chronically dissatisfied with research efforts in this field."<sup>7</sup>

There are no doubt many factors contributing to this state of affairs. Indeed, the sections of the thesis concerned with the evaluation of the Open University counselling services will describe some of the influences upon evaluation-type inquiries. More fundamentally, could it not be that a contributory cause of this all-round disillusionment is the "confusion" and "lack of agreed frames of reference"? If this is so, it would seem to be worthwhile to attempt to clarify this confusion, in the expectation that a clearer understanding of the nature of evaluation will help to improve the chances of carrying out effective evaluations.

In order to do this, it will be necessary to examine some of the positions which have been taken with respect to the nature of evaluation from within a philosophical framework. By doing this, it is possible to gain a number of fruitful insights into the subtleties of evaluative language and reasoning which provide an atmosphere for thinking about the nature of evaluation. Secondly, it will be necessary to delineate a model of the evaluation process. This model is intended to be general enough to encompass the widest possible range of specific evaluative situations and therefore of maximum utility. In one sense it represents a set of instructions for doing evaluation. In another, it is a representation of the processes and elements which are involved in many empirical evaluations. As a general model it applies not only to educational evaluation but to any evaluative context, although the thesis as a whole is embedded in an educational one.

It is probably wise to point out now that the notion of evaluation to be presented is one that the writer regards as closest to the basic meaning of the process of evaluating, that of making value judgements. As will be shown later, some of the current prescriptions which go under the name of educational evaluation are not, from the writer's point of view, strictly about evaluation. As Provus says, "Considerable misconception about evaluation has resulted from our failure to make clear distinctions between research and evaluation processes", which are "not to be equated".<sup>8</sup> In fact, much of what is called evaluation by practitioners is essentially a form of applied research, and much of the literature is primarily concerned with dealing with essentially research-oriented problems.

In the writer's view, research may often be undertaken as part of the cluster of activities associated with the business of evaluation, but the production of research results and findings does not in itself constitute the end product of evaluation. Instead, research and the provision of information - which after all is all that research does - is seen as but one component which is a necessary, but not sufficient, condition for the conduct of evaluation. It is with these other components that we shall be mainly concerned, and in particular with the notion of standards. The research aspect of evaluation seems to have received a good deal of attention from practitioners; it is the complementary and more strictly evaluative aspect of the subject which might lay claim to the title of 'strangely neglected topic' - at least within the field of educational evaluation.

This does not mean, of course, that the writer regards the contributions of current evaluation practitioners as wrong on their own terms, or of little value. The point is that to talk about problems of measurement, or of experimental design, or of research in natural settings, as if they were wholly problems of evaluation, is to confuse the issue. These are certainly problems associated with the execution of some kinds of formalised evaluation. But they also arise in contexts that are not normally regarded as evaluative. On the other hand, evaluation does have problems which are all its own. And it does involve activities and procedures which are separable from those of research, measurement and description. Moreover, it is his concern with these matters which seems likely to give the professional evaluator his distinctive role.

Educational evaluation is, therefore, a relatively new and important field as well as a somewhat confused and turbulent one. The aim of this thesis is to provide a modest contribution to this field by attempting to delineate a clearer view of the nature of evaluation.

The second part of the thesis focusses on a particular evaluation project, carried out in connection with the Open University's tuition and counselling services between 1971 and 1974. Although a good deal of time will be spent later in describing the project and its setting at the Open University, it will probably be helpful to give a general description of the University here. This is made particularly necessary because it is a rather unusual organisation.

The Open University is a unique institution in the British higher education system, having many features which distinguish it from conventional universities. The University received its Charter in 1969 and began its first teaching year in 1971. It is intended to cater for a geographically dispersed population of adult students, normally working in full-time occupations, and studying primarily at home. No formal qualifications are required for admission.

The University has no residential students and no centralised face-to-face teaching facilities. Courses are constructed at the University's headquarters at Milton Keynes, Buckinghamshire, by 'course-teams' consisting of academic, BBC and advisory staff. The courses are presented to the students by means of correspondence texts, together with television and radio broadcasts. A wide variety

Figure 1: Some Features of the Open University and its Organisation

Students

1. Adults, over the age of 21, dispersed throughout the U.K.
2. Study part-time from home
3. No formal qualifications required for entry

Courses

1. Generalist degrees, obtained by accumulating course credits
2. Produced by course teams at central HQ - academics, BBC staff, advisers, consultants, etc
3. Delivered as printed booklets, TV and radio broadcasts
4. Assessment by written work, computer-marked objective tests and examinations

Central Staff

1. Full-time academics attached to one of six Faculties
2. BBC staff, educational technologists and researchers
3. Administrative and technical staff: computer services, registry, etc

Regional Staff

1. Thirteen Regions, each with a Regional Office and Regional Director
2. Staff Tutors and Senior Counsellors supervise part-time staff
3. Part-time correspondence tutors, class tutors and counsellors

Study Centres

1. Rooms hired by the University, open each evening and in reach of most students
2. Meeting place for students, tutors and counsellors for tutorials, personal interviews etc
3. Often contain library of course materials, TV, radio, replay devices

of supplementary materials may also be used, such as laboratory equipment, sent to the students' homes through the post for 'home experiments', tape recordings, records, off-prints, slides and so on. Students are assessed on written assignments, end-of-year examinations and multiple-choice objective tests which can be marked by computer.

The courses are intended to be self-contained wholes so that a student can, in principle, complete the studies satisfactorily simply by working on the materials he receives through the post and the broadcast programmes. Although in some cases a short period of attendance at a 'summer school' is required before a student completes his course, the main emphasis of the teaching arrangements is on 'teaching-at-a-distance'; that is, teaching primarily by means which do not necessitate direct, face-to-face contacts between students and teachers. The designers of the University were, however, mindful of the fact that correspondence-type teaching institutions tend to suffer from high rates of non-completion by students and they therefore proposed to build into the University a safety-net, or 'fail-safe' system, which might offset such a tendency. To overcome the physical separation between the students and the 'teachers' who produced the courses at Milton Keynes, it was thought desirable that every student should have a representative of the University available locally. There would, in any case, need to be University staff on hand to mark students' written work.

There would, therefore, be at least one person in the system to whom each student would be linked for the duration of his course, and it was envisaged that these 'correspondence tutors', as they

were called, would be recruited on a part-time basis from other educational establishments. What was not known, however, was whether correspondence tutors would be geographically distributed in such a way as to make it practical for students to meet them in person. It was thought that a situation could arise whereby there would be a glut of suitable tutors in one part of the country and an insufficiency in another, so that some students would only be able to contact their tutor by post while others would be able to attend face-to-face sessions as well. It therefore seemed necessary to create another role which would meet the need to provide every student with a human link in an otherwise rather impersonal system. Recruits for this role, which was titled 'counsellor', would also be part-timers, and as they would not be discipline-based, as the as the correspondence tutors had to be, there would be a much better chance of being able to match them with groups of students whom it would be practical to meet on a face-to-face basis.

As a result of these considerations there arose what came to be known as the 'tuition and counselling system'. The way in which this system developed will be described in detail in later chapters, so only a sketch of the system as it was in 1971 will be given here. As well as receiving instruction in print and from broadcasts, students could use the services of three different types of person.<sup>9</sup> Correspondence tutors marked and commented on written work; class tutors gave face-to-face tutorial sessions; and counsellors conducted personal meetings with individual students as well as group sessions. These services were provided by part-time members of the University's

staff, who usually worked full-time in education elsewhere.

In 1971 more than 3,500 part-timers were employed.

Part-time staff were recruited on a local basis, and each member was attached to one of the University's administrative districts, called 'Regions'. There were twelve Regions in 1971<sup>10</sup> covering all of the United Kingdom, each one having a main Regional Office run by a Regional Director. Within each Region the University rented facilities from local colleges, schools and so on, which could be used for meetings of students with part-time staff. These 'study-centres' also contained television and radio receivers, copies of the printed course materials, and other supplementary matter.

The recruitment and supervision of the part-time staff was carried out by full-time members of the University called 'Staff Tutors' and 'Senior Counsellors'. The former were responsible for the correspondence tutors and class tutors, the latter for the counsellors. These 'supervisors' were based at Regional Offices.

Such was the Open University's tuition and counselling system in 1971. There have been various changes since then, but the overall structure has remained.

For a number of reasons, the tuition and counselling system became a focus of controversy in the early years of the University and partly as a result of this a research group was formed to conduct studies of its operations. The University's planners had envisaged that some kind of 'operational research' work would be needed to support its development,<sup>11</sup> and an Institute of Educational



Technology was formed in 1970. Although this was much more than a research institute, most of its members being attached to course-teams as advisers, it contained research groups looking at such things as the use of broadcasting, methods of structuring the printed materials, and students' study patterns. Research into tuition and counselling thus fell fairly naturally within the Institute's province, and a small group was set up to carry out the work. The writer was a member of the group from late 1970 until early 1974, and the evaluation of counselling was one of the main areas which concerned the group. This is the evaluation project, the evaluation of Open University counselling, that will be dealt with in detail here.

The purpose of including this case-study is two-fold. Firstly, it is intended to help to meet the demand for more case studies and accounts of the carrying out of research projects of this kind. As King says:

"Research on research, throughout the social sciences, is woefully neglected. Thus, while evaluation reports exist, case histories of the studies themselves do not ..." <sup>12</sup>

Similarly, Clark <sup>13</sup> has called for case studies in the closely related field of 'action research'. As a result of this kind of documentation it is envisaged that the problem common to such research projects might be identified. The rendering of this account is thus seen as valuable in itself as a source of data for the use of other researchers. It is also, perhaps, particularly interesting by virtue of the fact that the project was undertaken within one of the most unusual educational institutions in the world.

The second purpose of the case study is to provide a concrete empirical frame of reference to which the evaluation model can be applied. The result of this will not be to make an actual evaluation of the Open University counselling services, but rather to show why such an evaluation was, and is likely to be, a particularly problematic endeavour. Given that the evaluation model specifies certain conditions which must be met if an explicit evaluation is to take place, the problems of meeting these conditions will be illuminated by considering this particular case. The topic is a somewhat unusual one, insofar as educational evaluation has, perhaps, until recently been predominantly concerned with the evaluation of curricula and innovatory programmes. However, writers such as Stufflebeam have tried to extend the idea of evaluation to a much wider range of subjects connected with education, and certainly within the Open University evaluation has been seen as relevant to topics which are not strictly instructional. Indeed it might be said that once an evaluative philosophy takes root, there is virtually nothing which cannot be brought within its focus of attention. Whatever lessons there are to be learned from the examination of this particular case are therefore likely to be relevant to other educational contexts, if, as seems possible, the all-encompassing view of evaluation becomes increasingly applied within educational institutions.

On the other hand, it is true that the unusual nature of the Open University has created special problems for research and evaluation. This perhaps makes the topic of Open University

counselling even more of a special case than it might otherwise be. However, this fact makes it all the more relevant, given that the Open University can be seen as the first concrete manifestation of a revolution in educational organisation. For the present the Open University is almost unique, although its methods are beginning to be copied in other parts of the world. One day, Open University-type institutions may well be common, at least in the technologically advanced nations. Thus by considering problems posed by the Open University situation, insights may be achieved which may be applicable to similar kinds of institution which have been established, or which will be established in the future.

There is a further reason for using Open University counselling as a case study. As was noted above, it was as a member of a research group whose task was partly to evaluate Open University counselling, that the writer first became involved in the problems of evaluation. It was partly as a result of these experiences that the idea of clarifying the nature of evaluation arose. It is therefore, perhaps, natural to focus upon this particular case.

Writing as an insider does, of course, have both advantages and disadvantages. An advantage is that the insider has, by virtue of a direct and lengthy involvement in the situation, a more intimate understanding of 'what went on' than it would be possible for an outsider to obtain. Yet this personal involvement also brings dangers. The objectivity of the account may be doubted, given that in the course of his experience the writer was oriented

to the scene as an actor rather than as an observer. Moreover, as historians know, it is always wise to interpret documentary evidence with regard to the position, interests and motives of the authors. How can the reader be sure that the account is valid?

The fact is that he cannot be sure, no more than can the writer. Any report of events is always an interpretation, and it is as an interpretation that the account of the research into counselling is offered. As such it is left open to challenge. It is almost bound to be an oversimplification, but to convey the full details of the happenings of a period of three years would be both an impossibly lengthy and difficult task. No doubt it would also try the reader's patience too much. As an interpretation, then, the account embodies the selection of certain events, papers and conversations which seem to cohere around a theme, and which by their selection enable a story to be told which has some recognisable structure. The alternative would be a mass of unrelated material. The account has, however, been grounded in such recorded data as is available, and wherever possible references to documentary sources are given.

In an account of this kind it is inevitably necessary to relate mistakes and problems, arguments and dissensions, since, after all, no one is interested in the newspaper that only reports good news. However, it is not the purpose of the account to attribute praise or blame, nor to offer smug criticisms. This would be not only presumptuous, and perhaps unethical, but also out of place in a

thesis of this kind. Rather it is intended to provide an overview of the activities and circumstances associated with the project. Yet it is not enough simply to describe. A commentary has also been included which necessarily implies some sort of evaluative stance on the part of the writer. This seems inevitable if anything significant is to be said at all. As with all value judgements, the ones presented here are open to argument.

The general organisation of the rest of the material in the thesis is as follows.

In the next chapter, a historical account is given of the development of educational evaluation, with a view to showing that it has become a field of some importance. Following this, there is a summary of some of the views put forward by practitioners about the nature of educational evaluation and evaluation in general. These works are those which have been produced by those who are professionally engaged in conducting evaluative investigations. A number of confusions and disagreements about the field are then noted. Next, some preliminary consideration is given to the nature of evaluation and value judgements, and some distinctions between evaluation and description are drawn. A model of the evaluation process is proposed and an examination is made of some of its components. The nature of evaluative disputes is then considered and some methods are suggested which may be helpful in resolving them.

Chapter Seven provides the background to the Open University and surveys the development of the University's local tuition and counselling services with particular reference to the latter. This is followed by a summary of the activities and problems encountered by the Institute of Educational Technology's Tuition and Counselling Research Project. Finally, an examination is made of the problems involved in evaluating the counselling service and of some of the problems of integrating research and decision-making.

REFERENCES

1. A.B. Thomas, 'Success and Failure in Open University Counselling', Teaching At A Distance, 1 (November 1974).
2. C.J. Parsons, Theses and Project Work, London 1973, p.49.  
He is, of course, drawing on Kingsley Amis' Lucky Jim.
3. For example, S. Wiseman and D.A. Pidgeon, Curriculum Evaluation, Slough, Bucks, 1972.
4. See D.L. Stufflebeam, et al, Educational Evaluation and Decision Making, Itasca, Illinois 1971.
5. M. Parlett and D. Hamilton, Evaluation as Illumination: A New Approach to the Study of Innovative Programs. Occasional Paper No. 9, Centre for Research in the Educational Sciences, University of Edinburgh 1972.
6. Stufflebeam, Chapter One.
7. M. Guttentag, 'Models and Methods in Evaluation Research', Journal for the Theory of Social Behaviour, 1 (1971), p.75.
8. M.M. Provus, Discrepancy Evaluation for Educational Program Improvement and Assessment, Berkeley, Calif. 1971, p.161.
9. In fact one person could carry out all three roles, so that although about 3,500 persons were employed as part-time staff in 1971, they filled more than 5,000 posts.
10. Reorganisation later increased the number of Regions to thirteen.

11. See Department of Education and Science, The Open University: Report of the Planning Committee to the Secretary of State for Education and Science, London 1969, p.15.
12. C. King, 'Alternative Evaluation: The New Paradigm in Action', in D. Hamilton, et al., eds., 'Whatever Happened to Daniel Stufflebeam?', pre-publication draft of a work on educational evaluation, 1975, p.1.
13. P.A. Clark, Action Research and Organizational Change, London 1972, Chapter Thirteen.



## CHAPTER TWO

### THE RISE OF EVALUATION IN EDUCATION

It is probably true to say that fifty years ago educational evaluation was unheard of as a special field. Yet today it has its own literature, its own journals, its own practitioners, and its own history of applications. It is useful to try to see how the emergence of education evaluation as a problem network <sup>1</sup> in its own right has taken place, and to note some of the main features of its development. This is so partly because a good deal of the literature has been written by practitioners in the United States of America. The examination of the history of educational evaluation is largely set, therefore, in an American context, and the concerns expressed within the literature produced there can be better understood in relation to that context. It is also worth looking at something of the history of educational evaluation in order to gain some insight into the implications and significance which it has for the conduct of the educational enterprise at large. For in some formulations, educational evaluation can be seen to present a radical challenge to the assumptions and practice of 'traditional' education. And the application of an evaluative approach can sometimes lead to some rather bizarre results.

At this point we will not go deeply into the question of what educational evaluation is all about. A summary of practitioners'

definitions is held over until the next chapter, and the beginnings of a clarification of the concept until the chapter after that. Enough will be said, however, to give the reader a reasonable idea of the kind of activity being talked about.

### Educational Evaluation in the United States

In the American educational system, evaluation has become an institution. It is doubtful if more than a handful of people in Britain hold the job title 'evaluator', and it is only now that evaluation is becoming recognised as a specialism distinct from educational research in its other forms. In the United States, the situation is quite different. There one can find such organisations as the Center for the Study of Evaluation, the Evaluation Research Center, various Educational Research & Evaluation 'laboratories', and Evaluation Centers; journals concerned solely with evaluation, such as Evaluation Comment, which is published by the University of California Center for the Study of Evaluation; conferences devoted entirely to the subject of evaluation, such as the Conference on Four Case Studies in Formative Evaluation held at Bloomington, Indiana in July 1971; and perhaps most important of all, one will find Federal legislation which requires the evaluation of certain projects by law.

Most of these developments are of a fairly recent origin. In particular, the 1965 Elementary and Secondary Education Act gave a tremendous boost to the field of evaluation. But this Act can be seen as a yeast which leavened a mixture which had been prepared much earlier.

According to Stufflebeam,<sup>2</sup> one of the current definitions of evaluation is that evaluation is measurement, and in the context of education this usually means psychological measurement, or psychometrics. On this view, evaluation is a very old enterprise indeed. Merwin,<sup>3</sup> for instance, suggests that one of the world's first evaluators was the Emperor of China, who devised a proficiency test for his Imperial officials as early as 2200 BC. However, we can perhaps be forgiven for moving forward several thousand years, to the beginning of the Twentieth Century. For it was at this time that psychology emerged as a 'scientific' discipline separate from philosophy, and its practical implications and the drive to 'make education a science' enabled Ayres to remark in 1918:

"Knowledge is replacing opinion, and evidence is supplanting guesswork, in education as in every other field of human activity".<sup>4</sup>

In so far, then, as evaluation is regarded as to do with the application of the social sciences to education, the idea goes back at least half a century.

However, a generally acknowledged 'breakthrough' in educational evaluation came in the 1930's when Tyler gave it its "first and most enduring set of procedures".<sup>5</sup> His prescription for curriculum design and evaluation involved five main steps, namely:

1. Secure agreement on the aims of the curriculum
2. Express these aims as explicit learner behaviours or objectives

3. Devise and provide experiences that seem likely to enable learners to behave in the desired way
4. Assess the congruence of pupil performance and objectives
5. Vary the treatment until behaviour matches objectives.

The crucial notion was that of behavioural objectives. Once the outcomes of education were stated in behavioural terms, they would be accessible to measurement, and it would then be possible to see if the outcomes were in accordance with the objectives. Doing this was doing evaluation.

This model was, and is, very influential in the field of educational evaluation, although some rather different ones have emerged since Tyler's first exposition. But it was not until after the Second World War that the educational evaluation machine really gathered speed, and began to take on different forms. Tyler himself says that "Since World War II, and particularly in the last decade [the Sixties] profound changes have been taking place in educational evaluation",<sup>6</sup> and that "..... the field of educational evaluation is in a state of great ferment".<sup>7</sup> He accounts for this in terms of various historical and technical changes which have influenced American education, particularly at the secondary level.

One such development was the launching by the Russians of the first 'Sputnik' in 1957. This technological feat not only resulted in the channelling of millions of dollars into the American space effort, but also lead to a massive financial injection into the

education system for mathematics and science. One consequence of this was to create a demand among administrators, constructors and teachers for the evaluation of these new curricula. For the psychometricians, this meant a change of focus, away from a concern with the measurement of individuals and towards the evaluation of programmes.

Another factor which influenced the growth of evaluation is the creation of new technologies which have been applied in education. For example, developments in television and computers have generated new possibilities for the educational process. Since these technologies are often costly, and also represent a challenge to the status quo, there has been a natural demand for evidence to demonstrate that they actually work. At the same time, the widespread availability of computer facilities and the production of increasingly sophisticated techniques of statistical analysis has made Big Evaluation a practical possibility.

Perhaps the most important influence on evaluation in America has been the 1965 Elementary and Secondary Education Act. This Act provided one thousand million dollars for the use of schools which had a high proportion of disadvantaged pupils. It also required local school districts receiving these Federal funds to evaluate the effectiveness of the educational efforts which were supported by them. According to Provus, this was "the first educational legislation in history to call for local determination of the effectiveness of federally funded programs".<sup>8</sup> Many of the concerns about evaluation expressed in the largely American

literature stem from attempts to meet the evaluation requirement embodied in this legislation.

The wider concerns of the American public also seem to have contributed to the emergence of evaluation as a significant activity. It has become painfully obvious that American society is suffering from a budget of ills, and it is the education system which has been turned to as the saviour, and sometimes turned on as the culprit. Provus writes:

"America's public schools are presently being asked to redress the social ills of the nation: to reinvolve the disenchanted: to eliminate long-standing prejudice towards racial and minority groups: to select and train future professionals, skilled workers and menials (with little or no regard to initial aptitudes): to correct the effects of child-rearing abuses: to revitalize community agencies and service institutions: to re-educate parents in family-building and maintenance skills as well as participatory democracy: and to secure the continuous education of all Americans from infancy to grave. As a result, schools are wallowing in a flood of new programs pumped into existence at staggering cost. School boards are now asking about the effects of these new programs. Should they be replaced? Can better programs be found? Should more money be spent or less?"<sup>9</sup>

Perhaps the demand for reassurance that these problems are being solved has never been greater, hence the demand for evaluation.

Provus goes on to point out that this is a new situation:

"For a long time, the public schools of this country were recognised for the success of their work. Immigrants, the disenfranchised, the confused, the inept, and the outrageously boisterous, were more or less taught to read and write. That some never did learn was of little consequence to those responsible for running the schools. All that has changed. At first because of social awareness of education inequities associated with race and poverty, and more recently because of public aspiration for a more broadly based democracy, more responsive institutions, and greater self-determination, the schools have come under critical public scrutiny. Above all, the role of the federal government and our political system in shaping popular expectations and funding massive local efforts, has lead to a broad base of support for concern about the quality and benefits of this country's most pervasive and expensive enterprise." <sup>10</sup>

Pressure for evaluation has come, therefore, not only from the Federal Government and educational psychologists. It has also come from the American people at large. As Stufflebeam <sup>11</sup> points out, citizens have demanded evaluation to see if schools' funds are purchasing effective and efficient education. Whether this has been motivated by a desire to justify the withdrawal of those funds or by a desire to improve education is an open question.

It can be seen, then, that educational evaluation in America has grown as a result of developments in educational psychology and psychometrics; from the demands of the public and Government arising from changes in American society; and from the need to examine the

impact of new technologies on the education process. The 1965 Act can be regarded as instrumental to the institutionalisation of educational evaluation in the American educational system, and it is the consequences of this act for evaluation that will now be considered, together with some more recent developments.

#### Evaluation Under the Elementary and Secondary Education Act

The immediate effect of the Elementary and Secondary Education Act for evaluation was that it lead to many attempts to conduct evaluation research. But most school systems were unable to meet the demand for evaluation, and there were many problems in trying to meet the evaluation requirement.

"Title I was put into effect almost overnight with the intent of compensating for the limited experience of poor, deprived and handicapped children. Most districts had little or no time to plan their ESEA projects, much less to design project evaluations. Research staffs and facilities were inadequate .... basic information about students and program conditions were not available ..... Yet school personnel were told that evidence of project effectiveness, and reports summarising the evidence, were needed in order to ensure public and congressional support for the coming fiscal year."<sup>12</sup>

As a result, the Congress received descriptive reports, even though they constantly asked about program benefits.

---

The difficulties in implementing evaluation procedures were legion.



"'At last' we said 'curriculum evaluation has come into its own'. We began our work by oversimplifying the problem. It was, we decided, simply to find out whether new programs were better than the ones they replaced. We did not realize that our first problem was to find out what a new program was .... [We] discovered what educational researchers have known for a long time: When quasi-experimental designs are applied to the outcomes of new educational programs, generally no evidence of new program advantages over existing program is obtained.." <sup>13</sup>

Moreover, these difficulties were generally experienced:

"The history of recent public school programs is one of lack of documentation, lack of program control, and measured outcomes suggesting that there is greater variation within programs than between programs .... It is entirely possible that most public school evaluations are meaningless because they reflect the confusion of administrators over educational programs that are equally meaningless .... Today [1971] useful evaluation theory and practice are no better established in public schools than they ever were. However, the need persists." <sup>14</sup>

The evaluation requirement of the ESEA thus proved to be a difficult one to meet in practice, and the ESEA had three important effects on evaluation.

Firstly, as has been noted, many schools lacked the necessary research staff and facilities for evaluation research, and this helped to defeat the execution of evaluation studies. This created considerable problems, since the schools were often

unable to obtain ESEA funding because they could not meet the evaluation requirement. In response to this, specialised evaluation units were set up to service the schools' evaluation needs. This helped to establish a body of specialised evaluation practitioners, and thus to create a significant problem network for educational evaluation.

Secondly, the practical problems <sup>15</sup> experienced in the attempts to conduct evaluation investigations generated a quantity of literature dealing with the problems of evaluation research, as opposed simply to reports of particular evaluation studies. These materials constitute the bulk of the reference matter for those who operate within the educational evaluation problem network, providing its members with, as it were, a common culture.

Thirdly, the ESEA helped to associate the idea of evaluation with the movement for educational reform. Provus, for example, says:

"The evaluation clause of ESEA established evaluation as a necessary building block in the design of educational reform. The evaluation implications of ESEA could eventually have greater impact than the program itself". <sup>16</sup>

In other words, the ESEA is thought to have more significance in its potential for establishing the idea of evaluation as integral to the conduct of education, than in any of its direct outcomes in terms of the results of compensatory programs funded under its auspices. Evaluation, rather than just money, would be the main

catalyst of improvement and change. Evaluation, therefore, would become more and more important.

In summary, the ESEA helped to establish a body of professionals whose specialism is educational evaluation, a body of literature concerned with the problems of educational evaluation, and a body of practice and an awareness of educational evaluation in the educational system. It therefore represents an important landmark in the development of educational evaluation.

#### Evaluation and the 'Accountability Movement'

In order to appreciate some of the wider implications of the establishment of evaluation in education, it is worth seeing how it has become linked with accountability in American education.

According to Sciara & Jantz, "accountability could well become one of the most important educational movements in the decade of the 70's", and is "a formidable force in American education".<sup>17</sup> But what is accountability?

"Basically accountability means that public schools must prove that students at various levels meet some reasonable standard of achievement. The concept further implies that schools must show they use funds wisely - that expenditures justify educational outcomes."<sup>18</sup>

That is, accountability involves evaluation. But unlike evaluation under the ESEA, which was concerned with compensatory programs,

the notion of accountability can be applied to the school system as a whole. This wider application of evaluation was implied by President Nixon in his 1970 address on educational reform. There he said, somewhat ominously, that "We have, as a nation, too long avoided thinking of the productivity of the schools".<sup>19</sup>

Someone has said that evaluation almost always makes the 'client' look bad.<sup>20</sup> Certainly the consequences of 'looking bad' can be startling. With reference to accountability, Morris says:

"The crux of the issue, in greatly simplified form, seems to be that if teachers cannot teach pupils to read; solve mathematical problems; speak and write correctly; memorize principles, laws and formulas in science; and regurgitate names of people, places, things and dates found in social studies texts; then some business concern will guarantee to do so at a predetermined level of performance and cost."<sup>21</sup>

Indeed, that 'performance contracting' - the handing over of education to business - has been taken seriously in the United States is witnessed by the fact that President Nixon appropriated 6.5 million dollars from the budget of the U.S. Office of Economic Opportunity in 1971 for contracts with private industry. However, performance contracting is probably but one response to a diagnosis of failure.

Thus it is, perhaps, the accountability movement which has given evaluation an increasing impetus within American education - or perhaps it is the other way round. The need to hold teachers

and administrators 'to account' involves determining their 'success' and 'failure', in other words, evaluating them. Similarly, by evaluating, the possibility is opened up of using the results as a basis for attributing praise and blame. The relationship is symbiotic. Evaluation can thus be used as more than a means to improving the curriculum; it can also be used as a means to the end of controlling educators to a degree which is perhaps novel.

#### Summary of Developments in the United States

In summary, educational evaluation appears to have developed into a specialism in the United States as a result of a number of influences which have been brought to bear in recent times. As Rice, Buser & Ellis say:

" It appears that a new position is rapidly emerging in education, that of the evaluator. This position seems to be emerging in part from the demands of current education legislation, particularly the Elementary and Secondary Education Act; in part from awareness and appreciation for the significance of evaluation as a basic component of the decision-making process; in part from the demands of a society that is being asked to expend an over-increasing share of its wealth on education; in part from an appreciation of PERT [Programme Evaluation & Review Technique] by policy-makers; and finally by the measurement-evaluation specialists."<sup>22</sup>

Although initially focussed mainly on compensatory education program funded by the Federal Government, this focus now seems to have been widened via the accountability movement to cover the

'normal' aspects of education. With the growth of demand for evaluation there has come a growth in experience of conducting evaluative enquiries, and the emergence of a problem network whose focus is educational evaluation.

### Educational Evaluation in Britain

If educational evaluation in the United States is still in its infancy, it has hardly been conceived, by comparison, in Britain. The framework of evaluation institutions which has concretised the problem network in the United States is absent; there must be few, if any, who hold the post of 'evaluator'; and among the literature on evaluation there seem to be but a handful of works by British authors.<sup>23</sup> This does not mean, of course, that evaluative-type enquiries have not taken place,<sup>24</sup> but it does suggest that evaluation has not achieved the degree of prominence and significance that it has in America.

The reasons for this state of affairs may in part be illuminated by considering the development of educational research in general in Britain. According to Brehaut,<sup>25</sup> this development has been very slow.

Among the obstacles to the progress of educational research identified by Brehaut is the inherent distrust of the British of 'experts', particularly foreign experts. In the early part of the century, he says, many of the newest ideas in psychology (until recently, the King of the 'educational sciences') were formulated in Germany. The Americans readily took over these ideas at a time

when Britain was gripped with anti-German sentiments. Consequently, educational psychology developed more rapidly in America than in Britain. The British were also somewhat contemptuous of American educational research, with its transatlantic jargon and ethnocentric disregard of British work, and it is only "in recent years [that] the critical tone has become more moderate as British researchers look to America for leadership and, in some instances, for funds." <sup>26</sup>

At the same time, says Brehaut, the hostility of the academic community to the behavioural sciences helped to limit their growth in Britain. Departments of both psychology and sociology took a fairly long time to get established. There has also been a great deal of criticism of educational research, which Brehaut believes has had a largely negative effect.

If educational evaluation research has been inhibited by the same factors which have influenced educational research at large, we can nevertheless point to some recent developments in educational thinking which have helped to promote its growth in Britain. In particular, the emergence of educational technology from its audio-visual aids phase is likely to bring evaluation increasingly to the fore.

Educational technology is itself a term which has only recently acquired a position in the common language of education. According to Reid, <sup>27</sup> it was the establishment of the National Council for Educational Technology by the Department of Education and Science in 1967 which indicated that the term was acceptable

in official circles. The Council was to encourage research and development in education and training, in the armed forces, industry, schools and tertiary education. In the same year, the journal Programmed Learning added 'and Educational Technology' to its title. The editor said that programmed learning and audio-visual aids could be drawn together, the latter having been most closely associated with the term 'educational technology'.

Since then, educational technologists have been at pains to point out that educational technology is not just about audio-visual aids and programmed learning. Rather it is concerned with a systematic approach to teaching and learning. In a recent monograph on the nature of educational technology, Rowntree writes:

"Educational technology is not to be confused with electronic gadgetry ..... educational technology is as wide as education itself, it is concerned with the design and evaluation of curricula and learning experiences and with the problems of implementing and propagating them. Essentially it is a rational, problem-solving approach to education, a way of thinking sceptically and systematically about learning and teaching." <sup>28</sup>

He outlines a four-phase educational technology approach, involving objective-setting, design of learning, evaluation and improvement. Reid summarises the systematic approach of educational technology as having four aspects, specification of objectives, assessment of characteristics of students, specification of content, methods and materials, and assessment of effectiveness. Both authors include



evaluation as an integral feature of the educational technology approach, and Rowntree's remarks about the need for the kind of orientation he proposes suggest that some of the factors which have encouraged educational evaluation in America may be beginning to be felt in Britain:

"Nevertheless, we have all had to recognise that the sheer variety of new materials and media, and the enormous bulk of it, raise vexed questions of evaluation and choice for the teacher. This pressure from media ..... have [sic] allied with the pressures of curriculum reform, school population growth, youth unemployment, parental concern, student protest, employer anxiety, and the uncertainty of the future ahead of our students, to encourage us in renewed thinking about what education is and might become and what we should be doing about it." <sup>29</sup>

Perhaps one of the most important factors, however, which has led to the relatively inconspicuous presence of educational evaluation in Britain, is its different cultural milieu. Educational evaluation is often associated with a philosophy of self-conscious scepticism about the attainments and effectiveness of education. Yet, according to Brehaut, the British have lacked interest in efficiency, and have developed a 'cult of the amateur' in the administration and organisation of education. By way of contrast, professional administrators have long been the norm in the United States. Williams say of American society:

"Since systematic wealth-getting, technological achievement and productive organisation of effort

have been strongly sanctioned, pressure has been created to search for 'better methods', with the result that America epitomises high regard for efficiency in techniques. In this kind of social climate, there is a high sensitivity to such epithets as 'backward', 'inefficient', 'useless'." <sup>30</sup>

But in Britain, says Brehaut, "the whole history of the English educational system shows that efficiency has not been an important objective." <sup>31</sup> Similarly, Hamilton et al describe American culture as "essentially forward-looking, constructionist, optimistic and rational", and British culture as "largely backward-looking, conservationist, complacent and distrusting of rationality." <sup>32</sup>

Yet the introduction of new technologies into education; the development of novel institutions such as the Open University; the increasing pressures on educational finances; <sup>33</sup> and the spread of the evaluation-minded philosophy of educational technology, seem likely in the writer's view, to make educational evaluation an established practice in the British educational system.

### Summary and Comments

In this chapter we have given a brief overview of some developments in the field of educational evaluation in the United States and Britain. It seems that educational evaluation is a far more distinct activity in the United States than Britain, existing as a specialised field which is the province of professional practitioners. It has a history of applications, a body of literature, and a framework of institutions which cannot be found to a comparable degree in this country. To that extent,

educational evaluation is something new to the British educational scene. It is beyond the scope of this thesis to try to offer a comprehensive explanation of the relative lack of prominence of evaluation in the British educational system in comparison with the United States. We have already suggested that this is partly to do with the differential development of certain disciplines, especially psychology, in each country, differing social experiences of education and a different attitude to the running of educational enterprises. Instead, the admittedly limited knowledge we have of the history of education evaluation in these countries suggests some general features of the context of evaluative inquiries which are worthy of comment.

One of the distinctions which is sometimes made in the literature on evaluation, is that between informal and formal evaluation, and it is useful to introduce this distinction at this point. Discussions of the difference between the two are often introduced in an attempt to deal with the thorny problem of defining evaluation, and represent evaluation as a 'technical' activity significantly different from evaluation as an everyday activity. Practitioners tend to recognise that evaluating is a ubiquitous activity, and so find it necessary to distinguish the procedures which they advocate from their everyday counterparts. One way of doing this is to introduce the concept of 'formality'. Thus informal evaluation is characterised as being based on information derived from casual, unsystematic observation and involving intuitive, subjective judgements. Typically, this form of evaluation is also regarded as inadequate and undesirable. In

contrast, formal evaluation, the kind which professional evaluation practitioners seek to practice, is characterised as involving the application of systematic, objective, empirical methods of data-collection and, to a lesser extent, the use of explicit criteria of judgement.

When we talk about the problems of educational evaluation we are therefore primarily concerned with formal evaluation, although the basic components of this activity are, we will suggest, the same as those which are involved in everyday evaluative judgements. To talk about the history of educational evaluation is thus to focus on formal evaluation, since everyday evaluations of education must be as old as education itself.

The kinds of evaluative investigation which professional practitioners typically have in mind, are highly formalised. Stufflebeam's approach provides perhaps the most outstanding illustration of this, where a vast array of concepts and procedures are set out so as to encompass a large number of decision-making possibilities. Indeed the formalisation of the activity of evaluation goes hand in hand with its social institutionalisation, since the claim to a distinctive role as evaluator is dependent upon there being a distinct set of activities associated with it.

Formalised evaluation thus tends to be a highly specialised and technical activity which requires the application of significant quantities of resources, both human (professional evaluators, psychologists, educational technologists, etc.) and material (data-

processing facilities, etc.). Because formal evaluation tends to require this supporting structure of resources, it tends to be carried out in the context of large, formal organisations. On the one hand, it is large organisations which are capable of making big decisions. As more and more resources fall under the control of an organisation, the costs of making decisions, and of making mistakes, increase. At the same time, the scale of the organisation enables it to support a specialist technical staff to aid decision-making. On the other hand, the development of technologies of data-collection and processing have created new possibilities for organisational control, and have encouraged decision-makers to utilise the services of research specialists.

One of the immediate implications of the tendency for formal evaluation to be carried out in formal organisations is that there tends to be a division of function between the evaluators and the 'clients'. This division has been a source of problems for the conduct of evaluative research, and we will suggest that this aspect of the context of formal evaluation presents the evaluator with the important, and neglected, task of identifying clients' criteria of evaluation.

Another point is that because formal evaluation tends only to be supportable under the auspices of big organisations, be they government departments, universities, school systems or whatever, it tends to be applied to big problems. At the Open University, for example, evaluative questions arose about the effectiveness of whole sub-systems of the organisation, such as the tuition and

counselling system. There was some implicit agreement among researchers and clients that resources ought to be devoted to the study of highly significant problems, although there may have been disagreement on what those problems were. Similarly, in the United States we find evaluation being conducted of major national educational programmes such as Head Start. The result of this is to place very heavy demands on evaluative research inquiries, which they tend, in common with applied social research into major problems, not to have been able to meet.

If we were to attempt to characterise the role of formal evaluation in societies such as the United States and Britain, then we might say that it constitutes an attempt to increase the degree of control exercised by decision-makers. This is seen most clearly in the case of the 'accountability movement' where evaluation is the means to realising the goal of holding educators to account. It is, of course, possible to undertake evaluations, both formal and informal, without an orientation to action. Sometimes we may judge things simply for the joy of judging them. Formalised evaluations tend, however, to be conducted, at least 'officially', for the purposes of taking action. If an employee is inefficient, he will be fired; if this machine is efficient, we will invest in more of them; if this educational program is successful, we will repeat it; and so on. Formal evaluation could, in principle, be put at the service of anyone capable of either learning a minimal set of techniques himself or of hiring those who already possess them. In practice formal evaluation tends to be at

the command of institutional decision-makers in large organisations. It can thus be seen as intended to increase the controllability of such organisations. In so far as most organisations in education exist in a public domain, and cannot be directed in a dictatorial fashion by a single decision-maker without regard to statutory constraints and conventions, public opinion, and so on; and since formal evaluation tends to be introduced into major decision-making situations; then we tend to find not only that the evaluator is separated from the client, but also that he works with a number of clients whose interest impinge, directly or indirectly, on the decision at hand. The very existence of political institutions within organisations (committees, general meetings, Senates, Councils, etc.) acknowledges that conflict between parties is likely to arise. Formal evaluation might thus be seen to operate not only in the context of organisational control, but also in the context of organisational dispute. We shall argue later that calls for formal evaluation typically arise from a context of evaluation dispute, but we can simply note here that formal evaluation will tend to be influenced by a multiplicity of decision-makers, and that this potentially multiplies the problems of communication for the evaluator, and in particular the problem of identifying standards.

Another commonality which we find both in the experience of the United States and Britain is that formal evaluation seems to be called for when things seem to be going wrong. In America, for example, the Russians' success in the space race created a kind of

virility crisis, and more recently we find widespread concern over the effectiveness of the education system. For Britain, Rowntree lists a variety of anxieties; youth unemployment, student protest, parental concern, and so on. We do not seem to hear calls for formal evaluation when there is widespread satisfaction about a state of affairs. In a later chapter we will suggest some plausible reasons as to why this should be so, but whatever the explanation, the tendency for formal evaluations to be carried out in situations where things seem to be going wrong means that evaluation tends to be seen as an activity which will simply confirm everyone's worst fears. Given the organisational context of much evaluative research, if something is going wrong, someone is responsible for it, and as these people are often those whose activities are the subject of evaluation, this tends to make for difficulties in the execution of the evaluative inquiry. Examples of the kinds of problems that can arise will be given in our account of the evaluation of counselling at the Open University.

A final point we can consider is whether formal evaluation is likely to be a means to educational reform. As we have shown earlier, some practitioners have suggested that the introduction of formal evaluation of educational activities on a wide scale could do much to hasten educational reform and change. It may well be that if evaluative research demonstrates that a particular pedagogic technique is less effective than another, or that a given class size is associated with higher achievement than another, then the prospects for change in such matters will improve. It seems,



however, that experience to date is not encouraging, partly because of the inability of evaluative research to provide strong evidence, and partly because there is no guarantee that decision-makers will respond to research findings. Both of these are matters which will concern us later.

Much depends, however, on one's conception of evaluative research. If it is best conceived as an activity which determines the relation between a set of evaluative standards and some state of affairs, then it can do no more than just that. Evaluation research can be seen as oriented to discovering whether something is or is not as it should be; in giving its answer it does not say what to do about the situation. In that sense it is not oriented to creating or suggesting new technologies, though it may stimulate the search for them. On the other hand, to the extent that evaluative research provides information that is relevant to educational debate, it may help or hinder the particular groups who seek change.

In sum, the likely relation between evaluative research and reform in education is unclear, but it would perhaps be prudent not to expect too much from it.

We now move on to examine some practitioners' views of the nature of evaluative research.

REFERENCES

1. For problem networks see M.J. Mulkay, The Social Process of Innovation: A Study in the Sociology of Science, London 1972.
2. Stufflebeam, Chapter One.
3. J.C. Merwin, 'Historical Review of Changing Concepts of Evaluation', in R.W. Tyler, ed., Educational Evaluation: New Roles, New Means, Chicago, 1969.
4. Cited by Merwin, p.4.
5. Hamilton, 'The Objectives Model Revisited', p.2.
6. Tyler, p.1.
7. Ibid., p.4.
8. Provus, p.6.
9. Ibid., p.5.
10. Ibid., pp. 5-6.
11. Stufflebeam, p.268.
12. Provus, p.6.
13. Ibid., p.7.
14. Ibid., pp.7-8.

15. "The passage of the ESEA 1965, with its provision that projects funded under the act were to be evaluated, stimulated an unprecedented number of efforts to appraise educational programs. Most of these efforts were ill-designed, failing to recognise the complexity of the task and the lack of available instruments and personnel."  
R.W. Tyler, Introduction to Provus, pp.2-3.
16. Provus, p.9.
17. F.J. Sciara and R.K. Jantz, Accountability in American Education, Boston 1972.
18. Ibid.
19. Cited by G.D. Hottleman, 'Performance Contracting is a Hoax!', in Sciara and Hantz, p.259.
20. J. Davis, 'Great Books and Small Groups: An Informal History of a National Survey', in P.E. Hammond, ed., Sociologists at Work, New York 1964, p.217.
21. J.E. Morris, 'Accountability, Watchword for the 70's', in Sciara and Jantz.
22. D. Rice, R. Buser and J. Ellis, 'Educational Evaluation - A Model for Task Oriented Position Development', Contemporary Education, 41 (1970), reprinted in Sciara and Jantz.
23. The work by Wiseman and Pidgeon is a notable example.
24. For example, the evaluation of I.T.A. in the 1960's.

25. W. Brehaut, 'A Touch of Genius in Us All', Times Higher Education Supplement, 18 January 1974, pp.11-12.
26. Ibid., p.11.
27. R.L. Reid, 'Educational Technology: The Background and the Need', in D. Unwin, ed., Media and Methods: Instructional Technology in Higher Education, London 1969.
28. D.R. Rowntree, 'What is Educational Technology?', The Open University Institute of Educational Technology Monograph No. 1, p.1. First emphasis mine.
29. Ibid., p.2.
30. R.M. Williams, American Society, New York 1952, p.401.
31. Brehaut, p.11.
32. Hamilton, 'The Objectives Model Revisited', p.4.
33. University Vice Chancellor's have recently welcomed an enquiry into the relative cost-effectiveness of universities and other sectors of tertiary education.

### CHAPTER THREE

#### PRACTITIONERS' VIEWS OF EVALUATION

The rise of evaluation in education has naturally been accompanied by the emergence of a body of literature on the topic. Much of this emanates from the United States. In this chapter we will review some of the more recent of these works, with a view to identifying a spectrum of opinions on the nature of evaluation from the point of view of practitioners. Most of these works refer to specifically educational contexts, but we will also consider some general references to evaluation and evaluation research.

It will be appreciated that the length and complexity of some of the works makes it very difficult to give both a concise and accurate account of their contents. We will, however, focus on the central approach to evaluation of these authors so as to obtain a broad appreciation of their conceptions of the topic. Our main task is to describe the various ways of looking at evaluation, so that comments will generally be reserved until the end of the review.

N.E. Gronlund: Measurement and Evaluation in Teaching, 1968<sup>1</sup>

Gronlund begins by noting that some form of evaluation is inevitable in teaching. Already teachers use various sources of information to make day-to-day decisions. There is, however, some confusion about the meaning of evaluation in education. Sometimes evaluation is regarded as synonymous with measurement. A teacher who administers an achievement test may say he is measuring or evaluating achievement. On the other hand, qualitative descriptions of pupil behaviour are sometimes regarded as evaluation, as distinct from quantitative measurement. Gronlund's own definition of evaluation is "a systematic process of determining the extent to which educational objectives are achieved by pupils."<sup>2</sup> He emphasises that this enterprise is systematic, and that it demands a prior identification of educational objectives.

Evaluation, he says, is a more comprehensive and inclusive term than measurement. It involves both qualitative and quantitative descriptions of pupil behaviour, as well as value-judgements concerning the desirability of that behaviour. The main emphasis in evaluation is on the extent to which educational goals have been achieved. The main purpose of evaluation is to improve learning and instruction.

Behaviour to be evaluated includes the general categories of abilities and typical behaviour, and methods for obtaining information for evaluation include achievement tests, intelligence tests, interviews, questionnaires, anecdotal records and rating scales. Evaluation is not, however, a collection of methods for obtaining information about pupils; it is a process of determining the extent to which pupils achieve educational objectives.

M. Parlett and D. Hamilton: Evaluation as Illumination, 1972<sup>3</sup>

This paper is subtitled 'A New Approach to the Study of Innovatory Programs'. It aims to provide, or at least advocate, "a total reappraisal of the rationale and techniques of program evaluation",<sup>4</sup> although the approach suggested can be applied equally to traditional teaching as well as innovatory methods. The authors contrast the traditional experimental approach to evaluation with a 'social anthropology' approach.<sup>5</sup> Evaluation in terms of the experimental approach is regarded as an assessment of the effectiveness of an innovation by examining whether or not it has reached required standards on pre-specified criteria. The aims of illuminative evaluation are to study the ways in which programs operate, the influence of the school situation in which they are introduced, the way they affect students' intellectual tasks and academic experiences, and the advantages and disadvantages which are perceived by those most directly concerned. It aims to record the experience of participating in such programs, from both the point of view of teachers and students, and to identify the programs' most significant features. The main aim is description and interpretation rather than measurement and prediction.

Two important concepts lay at the heart of the illuminative approach. These are the instructional system and the learning milieu. Instructional systems are sets of teaching arrangements which are described in various formal documents. These formal blueprints are the ones normally used by evaluators to identify the objectives which are to be used in evaluation. However, enacted

instructional systems are rarely identical with the blueprints, and thus form an unilluminating basis for evaluation. The learning milieu is the social-psychological and material environment in which the students and teachers work together. This milieu is the context into which an innovation is placed, and it has a unique modifying influence upon it. Innovations cannot be regarded as closed, independent, discrete systems. Evaluation therefore needs to take into account both the instructional system, the learning milieu and the interactions between both.

Illuminative evaluation does not have a fixed methodology, but follows a general research strategy. The main stages are general acquaintance with the program-in-action, selection of significant areas for further enquiry, and description and explanation of significant areas. The main techniques used are observation, interviews, questionnaires and tests, and the collection of data from documentary and background sources. This approach thus encompasses the use of a much more wide-ranging set of data collection methods than the traditional one, and seeks to produce a rich and heterogeneous body of data.

The main purpose of evaluation studies is seen as to contribute to decision-making. The evaluator does not, however, make the decisions. He aims to provide an understanding of the program and to raise the sophistication of the debates surrounding decisions.



M. Provus: Discrepancy Evaluation, 1971<sup>6</sup>

Provus' approach is related directly to the evaluation of programs in American schools. His model for evaluation is intended to enable an estimate to be made of the effect of a program on students' performance. Interest in evaluation, he says, arises from a desire to provide the best possible educational programs and because of scepticism about existing programs. It is assumed that a problem exists, that a solution can be found and that evaluation will aid the solution process. Evaluation is seen as part of the problem-solving process, and aims to support decision-making.

The model is intended to be used with 'instant installation' programs in American schools, as opposed to 'canned programs'. The former are hastily conceived efforts mounted within and by school districts; the latter are properly prepared and tested programs produced by specialists.

Five stages of evaluation are proposed:

1. Design
2. Installation
3. Process
4. Product
5. Cost

At each stage a comparison is made between the program as it exists and a standard. Where the comparison reveals a difference, this is called a 'discrepancy'. The evaluator begins by establishing

the design of the program. This is done by asking those involved about their objectives, the need for teacher-skills, materials and so on. In this way the features of the school-designed program are made explicit.

The design is evaluated in terms of its theoretical and structural soundness. Experts are asked to examine the design and comment on theoretical inadequacies, such as that the design assumes a higher average reading speed among the target population than is warranted. The structural evaluation takes place with reference to a set of 'generalised design criteria'. Seeing the program in terms of Inputs, Processes and Outputs, the design criteria require that relevant variables with appropriate values be specified for each of these stages. This design stage helps to show up weaknesses in the initial design, and identifies or creates standards for the evaluation of the processes and products.

At stage 2, the evaluator examines program activities to see if the participants are carrying out the tasks specified by the design.

At stage 3, interim products are identified and their relation to processes is determined, the interim products being those which must be achieved en route to major goals. The processes are regarded as acceptable if they have produced the interim products intended; if they have not, the evaluator returns to stage 1.

At stage 4, the task is to see if the combined interim products produce the major product or goal. Probus is undecided as to how

this is to be done, but expects it to involve the use of experimental or quasi-experimental research designs.

Stage 5 involves a cost-benefit analysis of the program in relation to comparable programs.

For Provus, the aim of evaluation is not simply to measure outcomes. It must also examine every aspect of design, installation, operation and cost. It is particularly important, says Provus, for the evaluator to document the implementation of a program, since this is one important area which is neglected in traditional approaches. As he points out, there is little value in knowing that one has arrived at the wrong destination if one does not know the route used to get there. If the design is sound, then an important source of 'error' is implementation of the design. In general Provus sees that evaluation must do more than exhibit results; it should also try to explain why the results occur, something which he believes current evaluations have largely failed to do. This, we can note, implies a much broader task for evaluation than the traditional objectives approach normally allows.

M. Scriven: The Methodology of Evaluation, 1967<sup>7</sup>

Scriven begins by noting that current conceptions of educational evaluation are unsatisfactory in both practical and philosophical terms. Evaluation has both goals and roles. The goals of evaluation are to answer questions about educational instruments (personnel, procedures, programs, etc.) such as what it is they do, whether they do it well, whether they do it better

than something else, and whether they are worth the cost. The process of evaluation is basically similar whatever we are evaluating. It involves gathering data on performance and combining them with weighted sets of goal scales, so as to obtain numerical or comparative ratings. Also included is the justification of the instruments used to collect the data, the weightings and the selection of the goals. Scriven thus introduces the important idea that evaluation involves justification.

Evaluation has not only goals but also roles. There are many different roles for evaluation, ranging from decisions about purchases of educational materials to decisions on reward and punishment in the factory, prison or school. Given its role in a particular setting, evaluation may then be oriented to two important types of end. In one case the aim is to help improve a developing product, which is the province of formative evaluation. In another, the aim is to inform decisions on whether to adopt a finished program, and this involves summative evaluation. Although evaluation may be undertaken in different ways, it always involves estimations of worth, value or merit.

Scriven recognises that a good deal of the unease about evaluation arises from the feeling that value judgements are subjective and unscientific. However, he believes that can usually be brought into the realm of rational debate, and that the difficulties which surround moral and non-moral judgements, in the context of curricular evaluation, should not be used as an excuse for avoiding a systematic approach to evaluation.

Scriven goes on to make a further important point which carries the idea of evaluation beyond the traditional performance-against-objectives model. He says that determining the extent to which a program meets its objectives is in itself insufficient, since this makes no account of the worth of the objectives themselves. Evaluation should thus encompass both the evaluation of performance and the evaluation of goals.

In the previous section, we noted that Provus saw the explanation of results as an important part of evaluation. Scriven touches on a similar point by saying that there is a tendency in the literature for evaluation to be seen as to do with interpretation or explanation. He points out, however, that it is preferable to distinguish explanation and evaluation, since data of a different kind will be required for each.

D.L. Stufflebeam, et al: Educational Evaluation and Decision-Making, 1971<sup>8</sup>

This work represents perhaps one of the most comprehensive considerations of evaluation seen from the point of view of its role in decision-making. Evaluation is defined as "the process of delineating, obtaining and providing useful information for judging decision alternatives."<sup>9</sup> The evaluator's main job is to provide information for decision-makers. Evaluation is thus a process which is divided between both evaluators and decision-makers. Only the decision-maker can specify his informational requirements, yet he needs the evaluator to generate this information for him. The

evaluator does not actually take the decisions about educational policy, but he does service them with the object of improving the rationality of the decisions.

Evaluation can be seen as having a general monitoring function in relation to educational activities, and a specific problem-solving function in relation to decisions to try to effect change. The evaluator must work with the decision-maker to identify areas requiring change, and then select the most appropriate problem-solving strategy. This strategy will vary according to the magnitude of the proposed change, and the amount of information available on how the change is to be brought about. Information must be provided about alternative strategies for change, the problems involved in and the extent of their implementation, and the results of their implementation.

Evaluation is thus seen as a process which is integral to the conduct of the educational enterprise. It must be eclectic in its methodology. The experimental method is inappropriate to evaluation research, which must operate in a natural setting rather than the laboratory. The requirements of the experimental method impose conditions which usually cannot be met, and seldom provides positive results. Evaluation must meet not only a number of scientific criteria of adequacy, but also a number of practical criteria. The evaluator must use the best means available to him to collect information, and not impose conditions which cannot be fulfilled. The aim of evaluation should be to improve rather than prove.

E.A. Suchman: Evaluative Research, 1967<sup>10</sup>

Evaluation, says Suchman, is basically an appraisal of value, a judgement of worth. The process of evaluating is highly complex and subjective. The task of scientific evaluation is to control this intrinsic subjectivity. Evaluative research has no special methodology, and must be judged by accepted scientific standards. One of the main difficulties in conducting such research is the need to reconcile scientific needs with administrative ones. This problem is rarely present in pure research.

There is confusion over the meaning of the term 'evaluation'. Terms such as 'assessment', 'judgement' and 'appraisal' are used interchangeably with the term 'evaluation'. Thus the American Public Health Association's 'Glossary of Administrative Terms in Public Health' defines evaluation as:

"The process of determining the value or amount of success in achieving a predetermined objective.

It includes at least the following steps:

Formulation of the objective, identification of the proper criteria to be used in measuring success, determination and explanation of the degree of success, recommendations for further program activity."<sup>11</sup>

Thus inherent in evaluation is the process of assigning value to some objective and then determining the degree of success in attaining the objective. Riecken recognises these two ideas and defines evaluation as "the measurement of desirable and undesirable consequences of an action that has been taken in order to forward

some goal that we value."<sup>12</sup> Riecken adds the notions that the object of study in evaluation is an activity, and that the activity may have negative as well as positive consequences. Evaluation is the process whereby man checks on his ability to influence other men and his environment. Often the activities to be evaluated are directed at planned social change.

Hyman also relates evaluation closely to the notion of planned social change. His definition is that evaluation is "the procedures of fact-finding about the results of planned social action."<sup>13</sup> For Hyman, evaluative research is applied research intending to study the effectiveness of the application of basic knowledge. Evaluation therefore "connotes scientific method, but has characteristics which distinguish it from that type of research whose objective is the accumulation and analysis of data in order to formulate hypotheses and theory for the sake of new knowledge itself, irrespective of judgment of the value of the knowledge."<sup>14</sup>

For others, effectiveness is the key term in the definition of evaluation. Greenberg and Mattinson<sup>15</sup> see evaluation as a measurement of the effectiveness of a program in terms of the program's objectives. James simply says that evaluation is "the measurement of success in reaching a stated objective."<sup>16</sup> Anderson, however, regards evaluation as "measuring achievement of progress towards predetermined goals", and as "concerned with determining whether the goals themselves are valid."<sup>17</sup>



Klineberg defines evaluation as "a process which enables the administrator to describe the effects of his program, and thereby to make progressive adjustments in order to reach his goals more effectively."<sup>18</sup> He also regards one of the purposes of an evaluation study as to determine reasons for success and failure, thus going beyond the measuring of effects to the discovery of why certain effects are produced and how they can be brought about.

Having summarised these views, Suchman recognises that evaluation studies may take many forms. He makes the distinction between evaluation and evaluative research. Evaluation is the general process in which the worthwhileness of an activity is judged, regardless of the method employed. Evaluative research, however, is the use of scientific method in order to make an evaluation. Evaluation is thus "the determination (whether based on opinions, records, subjective or objective data) of the results, (whether desirable or undesirable, transient or permanent, immediate or delayed) attained by some activity (whether a program or part of a program, a drug or a therapy, an on-going or one-shot approach) designed to accomplish some valued goal or objective (whether ultimate, intermediate or immediate, effort or performance, long or short range)."<sup>19</sup> Evaluative research is one method, but a most important one, of doing evaluation.

C.R. Wright: Evaluation Research<sup>20</sup>

Evaluation research aims to determine the relative success or failure of some programme of activities. The application of social science techniques to the appraisal of social action programs

has come to be called evaluation research. Evaluation research can be distinguished from other kinds of research by its purpose and the conditions under which it is conducted. The purpose of evaluation research is "to provide objective, systematic and comprehensive evidence on the degree to which a program achieves its intended objectives."<sup>21</sup> It is also concerned with identifying unintended consequences.

Evaluation research is unlike either exploratory, explanatory or descriptive research. Exploratory research aims to test theoretically significant hypotheses. Descriptive research aims to document the existence of certain conditions at a given moment or over time. Although good evaluation research may seek explanations of a program's success or failure, its main concern is to obtain basic evidence on effectiveness.

The special conditions under which evaluation research is conducted is that an objective, outside investigator serves a client.

Evaluation research may aim to compare the effectiveness of different programs, replicate prior evaluations of the same program or study the long-term effects of some program.

The basic procedure of evaluation research is:

1. Identify objectives of the program and possible unintended outcomes. These must be translated into measurable terms

2. Construct a research design which should ideally be that of the controlled experiment
3. Develop and apply the research procedure
4. Devise indices of effectiveness which relate measures of effect to expectations of effect
5. In the final stages of the evaluation, try to explain and understand how the effects are produced

The main goal of evaluation research is to measure the effectiveness of a program objectively.

R.E. Stake: The Countenance of Educational Evaluation, 1967<sup>22</sup>

Stake begins by drawing the distinction between formal and informal evaluation. Informal evaluation depends on casual observation, implicit goals, intuitive norms and subjective judgement. Formal evaluation, on the other hand, depends on checklists, structured examinations by peers, controlled comparisons and standardised tests of students. Usually, says Stake, informal evaluation is the rule, but even when formal evaluations have been carried out there has been dissatisfaction. They have often proved costly and have tended to produce results which are of little relevance to educators' concerns.

The basic characteristics of evaluation are the evaluation acts, the data sources, the congruences and contingencies, the standards and the uses of evaluation. The two basic acts of evaluation are description and judgement. An educational program

can only be fully understood when it is fully described and fully judged. Most evaluators, however, have chosen not to judge. They often feel incapable of doing this, and suspect that if they did so they would cease to obtain support for their work and would be denied access to data. Whether evaluators are in the best position to pass judgement is, for Stake, an open question. Nevertheless, he sees the processing of the judgements of others as an increasingly important task for the evaluators. Such judgement data can be obtained from groups such as students, parents, teachers and subject-matter experts.

In describing a program, the evaluator should include three types of data. Firstly, he should describe antecedents, that is conditions existing prior to teaching and learning which may have some effect upon outcomes. Secondly, he should describe transactions, the sequences of interactions between the educational actors. Thirdly, he should describe outcomes, which are the achievements, skills, attitudes, and aspirations of students which results from the educational experience. It might also include its effects on teachers, administrators, and so on, as well as on physical resources, such as the depletion of equipment. Long-term as well as short-term effects might be included.

This descriptive data can be divided into intents and observations. The evaluator must gather statements of intent concerning the program as a whole, including both the student behaviour that is intended and the teaching behaviour that is intended as well as the antecedent conditions that are expected.

The description of what actually happens proceeds largely by way of observation, either of an immediate kind or through the use of special instruments; interviews, tests, questionnaires and so on.

The intents are evaluated in terms of their logical contingency. Is there any logical relation, for example, between the intended transaction and the intended outcome? Observations are evaluated in terms of empirical contingency. For instance, does that transaction really result in that outcome? Finally, what happens is evaluated in respect of what was intended, that is in terms of congruency. The overall rationale of the program also gives grounds for evaluating intents, so that it can be asked whether the intents can be derived from the overall rationale. The results of these activities can themselves be subjected to the evaluation of participants, experts and so on, to determine, for example, how much incongruency is acceptable.

Stake now says that there are two approaches to the evaluation of a program. One is on the basis of absolute standards, and the other is by reference to the characteristics of other programs. There are usually many different sets of absolute standards, since there tend to be numerous reference groups and points of view. Judging for Stake, however, is not the comparison of reality with the standards but the business of choosing which standards to heed. From both the application of absolute standards and those derived from the examination of other programs, a composite rating of merit can be obtained (although it is not clear how) allowing recommendations relevant to educational decisions.

Table 11 Summary of Practitioners' Views

AUTHOR(S)	DEFINITION	PURPOSES	METHODS
GRONLUND	A systematic process of determining the extent to which educational objectives are achieved by pupils.	Improvement of learning and instruction.	Achievement tests, intelligence tests, interviews, questionnaires, anecdotal records, rating scales.
PARLETT and HAMILTON	The process of describing and interpreting enacted innovative programs.	To illuminate complex processes occurring in enacted programs and improve decision-makers understanding of them.	Eclectic approach. May use observation, interviews, questionnaires, check-lists, tests, documentary sources. Formal experiments inappropriate.
PROVUS	The comparison of the reality of a program with some standard with a view to identifying discrepancies.	To improve programs by providing data relevant to program modification and to determine the effects of programs on students.	Interviews, questionnaires, tests, panel methods, cost-benefit analysis, experimental or quasi-experimental designs.
SCRIVEN	Gathering performance data and combining them with weighted sets of goal scales to yield numerical or comparative ratings.	Can be used for many purposes, but is broadly to establish and justify the merit or worth of something.	Emphasis on experimental methods.
STUFFLEBEAM et al	The process of delineating, obtaining and providing useful information for judging decision alternatives.	To provide information to decision-makers and help improve the rationality of educational decisions.	Vary with change-settings and types of decision. Experimental methods only partially relevant.
SUCHMAN	The general process in which the worthwhileness of an activity is judged.	Improve control over planned social change.	Scientific evaluation requires the use of scientific method. Central method must be experiment.
WRIGHT	The process of determining the relative success or failure of some program of activities.	To provide objective, systematic, comprehensive evidence of the intended and unintended effects of programs.	Ideally use experimental methods.
STAKE	The activity of describing and judging educational programs.	To aid the development of programs and to inform decisions concerning the selection of programs.	Combine observational data and judgement data.

Comments on Practitioners' Views

From this brief summary of some recent works on evaluation, it can be suggested that a variety of different emphases have been placed on the term. As long ago as 1955, Klineberg was led to remark that "although this is perhaps not the place to enter into an argument over definitions, it would certainly be helpful if the term evaluation were not used quite so indiscriminately as it has been in the past."<sup>23</sup> Some years later, whilst considering the problems of evaluating public health programs, Fleck said:

"Unfortunately, the different groups interested in having public health programs evaluated in New York State, did not provide a definition of evaluation as a process. Indeed, many program directors regard the concept as nebulous, impractical and as sort of an investigation."<sup>24</sup>

Similarly, Gronlund, writing in 1968, noted that "there is some confusion concerning the meaning of the term evaluation as it applies especially to education."<sup>25</sup>

More recently, Carter writes that "while most applied social scientists agree on the need for program evaluation, in practice, there is little consensus as to the characteristics of such research."<sup>26</sup>

That the question of definitions is important is witnessed by the fact that writers often take some trouble to clarify what might be meant by evaluation. Stufflebeam,<sup>27</sup> for instance, considers the problem of definition as one of the causes of the 'sickness' of evaluation. He goes on to outline three definitions which have

been widely held, namely that evaluation is largely a measurement activity, that evaluation is the determination of the relation of performance to objectives, or that evaluation is the activity of professional judgement. He concludes that all these definitions have serious disadvantages. The first has led to too much concern with the design of measuring instruments, and has disassociated evaluation from the idea of value judgement. The second has led to too much concentration on the problem of operationalising objectives, and too much concern with the end product of achievement. The third de-emphasises the need for objective data collection. Provus similarly considers possible definitions of evaluation in relation to the evaluation of educational programs.

This uncertainty about the proper use of the term evaluation is reflected in the attempts to redefine the field. Stufflebeam explicitly recognises the terminological difficulty with his own definition:

"Because it is so different from current definitions, it may seem that the proposed definition of evaluation based on the decision-making process should have a name other than evaluation. Putting a new label on this process could possibly prevent many misunderstandings, misinterpretations and false imputations between what is meant here and what is always understood by the term."<sup>28</sup>

However, he finally decides that the use of the term evaluation is justified and that his definition is close to the root meaning of the term. Scriven,<sup>29</sup> however, in a critique of Stufflebeam's



model, says that some types of seemingly evaluative activity have been excluded by Stufflebeam's definition, and that some types of seemingly non-evaluative activity have been included under the head of evaluation. He also described the model put forward by Stufflebeam as "about the most complicated and confusing way of analysing the practical procedures of evaluation that I can imagine, and certainly the most complicated one I have ever seen."<sup>30</sup> He advocates a radical way out of the jungle of evaluative distinctions:

"The less jargon we can get by with the better; let's junk 'formative' and 'summative' and all these other terms, 'instrumental' and 'consequential' and so on, along with funny terms like 'context evaluation', and let's see if we can produce equally good evaluators in less time without them, or better evaluators in the same time".<sup>31</sup>

Parlett's attempt to develop new approaches to evaluation have led to a similar problem over terminology. Thus having proposed a new 'paradigm' to replace the traditional experimental approach, he remarks that "the type of study proposed here is, perhaps, so different that a new word is necessary",<sup>32</sup> 'evaluation' seeming to have the wrong connotations. He seems to have eventually settled for the term 'illuminative evaluation'. But in a later work in which some specific examples of the illuminative evaluation approach are given, some doubt is expressed as to whether they might properly be called 'evaluations' at all.<sup>33</sup>

It is, perhaps, fair to say that there is some confusion and disagreement over what evaluation is supposed to be about.<sup>34</sup> At the moment, it is hard to see what is particularly special about

evaluation and evaluative research. In terms of what is done, it sometimes seems to be little different from such activities as 'action research', measurement or simply data-collection. On the other hand, mention is often made of the need to establish criteria, the determination of success or effectiveness and so on, that is the idea of value judgement. It seems that many 'evaluators' have so conceived of their task as to ignore some of the central ideas which seem to be central to the concept of evaluation, or at least minimise their importance. Thus some of the issues which do seem to be inherent to the concept of evaluation have, perhaps, received less attention than they deserve. It might thus be regarded as useful to undertake an examination of the general idea of evaluation with a view to highlighting some of the more important aspects of the process. To this rather basic issue we turn our attention in the next chapter.

REFERENCES

1. N.E. Gronlund, Measurement and Evaluation in Teaching, New York 1968.
2. Ibid., p.6.
3. M. Parlett and D. Hamilton, Evaluation as Illumination: A New Approach to the Study of Innovatory Programs, Occasional Paper No. 9, Centre for Research in the Educational Sciences, University of Edinburgh, 1972.
4. Ibid., p.1.
5. Provus appears to have reached similar conclusions about the value of the 'agricultural-botany' approach. He writes: "It appears that the descriptive methods of the historian or anthropologist and the case-study method of the psychiatrist and sociologist are more appropriate to the task of educational program evaluation than the experimental methods of the psychologist or biologist". Provus, pp.iii-iv.
6. M.M. Provus, Discrepancy Evaluation for Educational Program Improvement and Assessment, Berkeley, Calif. 1971.
7. M. Scriven, 'The Methodology of Evaluation', in R.W. Tyler, R.M. Gagné, and M. Scriven, Perspectives of Curriculum Evaluation, AERA Monograph Series on Curriculum Evaluation, No. 1, Chicago 1967.
8. D.L. Stufflebeam, et al, Educational Evaluation and Decision-Making, Itasca, Illinois, 1971.
9. Ibid., p.40.

10. E.A. Suchman, Evaluative Research: Principles and Practice in Public Service and Social Action Programs.  
New York 1967.
11. Ibid., p.28.
12. Ibid., p.28.
13. Ibid., p.29.
14. Ibid., p.29.
15. Ibid., p.29.
16. Ibid., p.29.
17. Ibid., p.29.
18. Ibid., p.30.
19. Ibid., pp. 31-32.
20. C.R. Wright, 'Evaluation Research, in D.L. Sills, ed.,  
International Encyclopaedia of the Social Sciences,  
17 vols., 1968.
21. Ibid., p.198.
22. R.E. Stake, 'The Countenance of Educational Evaluation',  
Teachers College Record, 68 (1967).
23. O. Klineberg, 'The Problem of Evaluation', International Social Science Bulletin, 7 (1955), p.347.

24. A.C. Fleck jnr., 'Evaluation As a Logical Process',  
Canadian Journal of Public Health, 52 (1961), p.185.
25. Gronlund, p.6.
26. R.K. Carter, 'Clients' Resistance to Negative Findings and  
the Latent Conservative Function of Evaluation Studies',  
American Sociologist, 6 (1971), p.118.
27. Stufflebeam, Chapter One.
28. Ibid., p.44.
29. M. Scriven, 'Educational Evaluation: Noble Profession,  
Pedestrian Practice', paper presented to an AERA Symposium,  
February 1971.
30. Ibid., p.35.
31. Ibid., p.36.
32. M. Parlett, 'Evaluating Innovations in Teaching', in H.J. Butcher  
and E. Rudd, eds., Contemporary Problems in Higher Education,  
London 1972, p.153.
33. Thus in Hamilton the introduction to the first example of a study  
conducted under an illuminative approach says that "clearly it is  
debatable whether any of the shadow studies is an evaluation  
report at all". (p.6).

34. Not all practitioners would agree on the magnitude of this. Taylor, De Corte and Swinnen say: "There are almost as many idiosyncratic evaluation models that are currently available as there are people who write them. The idiosyncracies notwithstanding, it would be difficult to discern conceptual differences." P.A. Taylor, E. De Corte and K. Swinnen, 'Standards for Judging Instructional Effectiveness: A Problem in Educational Evaluation', Manitoba Journal of Education, (1970), p.5.

## CHAPTER FOUR

### THE NATURE OF EVALUATION

The last chapter revealed that there are a number of different definitions of educational evaluation currently available, and tried to establish that there is some acknowledged confusion about the nature of the topic. In particular, it was suggested that the idea of valuing in relation to the evaluative enterprise has received relatively little attention from evaluation practitioners, and it is to the fundamental idea of evaluation that attention is now turned.

The need for a clarification of the concept of evaluation is based upon the view that current definitions show an imbalance of emphasis among the concepts related to evaluation, and in some cases violate the established meaning of concept altogether.<sup>1</sup> The result of this is to leave evaluation with an insufficiently clear sense of its own identity, and perhaps more importantly, may lead to misconceptions and misunderstandings with the clients of evaluative research.

The purpose here is therefore to clarify the concept of evaluation, in the belief that clarity of thinking is a pre-requisite for controlled and effective action. In doing this, one is necessarily concerned with examining the language of

evaluation, and it might be said that it is futile to attempt to establish the meaning of evaluation, since anyone is free to use and define their terms, particularly in 'scientific' disciplines, in any way they choose. However, it is not intended that the meaning, or the definition, should be discovered. Rather it is hoped that the definition of evaluation given will be an adequate, useful and justifiable one and thus have some claim to being an acceptable one. At the same time, it is intended that the common-sense connotations of the term 'evaluation', which give the term its significance in everyday thought, will be retained in the final formulation. For it is surely not the case that anyone is free to define his terms as he wishes, unless he wishes to be regarded as a kind of conceptual anarchist. For, as Thouless says:

"If our thought is to be clear and we are to succeed in communicating it to other people, we must have some method of fixing the meaning of the words we use."<sup>2</sup>

Assuming therefore, that the term 'evaluation' does serve to reflect some real distinction among human phenomena (an assumption, of course, which may be challenged), the task is to make clear the implications which are but vaguely and partially formulated in everyday thinking.

But how are criticisms of the definitions offered for a term to be justified? What does it mean for a definition to be illegitimate? Thouless provides the general criterion that a definition is inadequate if it fails to serve the purpose of clear



communication, and it is contended here that the term 'evaluation', as currently used by evaluation practitioners, is just such an inadequately defined word. More specifically, Thouless suggests that inadequacy arises either from vagueness or from the infringement of the boundaries of other concepts. Thus to say that evaluation is "the process of determining the value or worth of a process or product"<sup>3</sup> is to be decidedly vague. It is certainly difficult not to be vague when considering evaluation, but it is hoped that the analysis here will give the concept a considerable degree of substance.

Infringement of boundaries can occur in two ways. One is by giving the concept an extension which is already that of some other word. Thus to call an activity which is the collection of data for use when comparing an existing state of system with a desired one, a type of evaluation, is to commit this error.<sup>4</sup> It is really part of the extension of concepts such as 'research', 'applied research' or 'data-collecting'. To subsume this activity under the concept 'evaluation', is to rob the latter of its value as a distinction. The other method of infringement is to equate a term with only a part of its established extension. Thus to only allow Christianity as a member of the extension of the concept 'religion' is to commit this error. In the case of evaluation, the definition given above - "the process of determining the value of a process or product" - ignores the fact that many phenomena which are neither processes nor products can be subjects of evaluation, such as imaginary, fictional and supernatural 'entities', (e.g. 'God').<sup>5</sup>

Of course, existing writers on education evaluation have all contributed something to our understanding of the idea of evaluation. All provide some glimpse of what evaluation might be about. Yet one feels they have only partially unravelled the concept. A similar impression is gained from the works on the philosophy of ethics; every 'school' seems to have a point, but none to have got all the points together in one place! Perhaps by bringing together the parts of the jigsaw, a new picture will emerge in which the parts take on a fresh meaning in relation to the whole.

The following quotation from Lavoisier will serve as the finale to this introduction:

"We cannot improve the language of any science without at the same time improving the science itself; neither can we, on the other hand, improve a science without improving the language or nomenclature which belongs to it. However certain the facts of any science may be, and however just the ideas we may have formed of these facts, we can only communicate false impressions to others while we want words by which these may be properly expressed."<sup>6</sup>

#### A Philosophical Consideration of Evaluation

To undertake an examination of the nature of the concept of evaluation is to suggest an essentially philosophical approach. Perhaps it would be desirable to regard the present investigation as simply a thoughtful examination of the concept of evaluation, so as to free it from the connotations of philosophical analysis

as a never-ending analysis of ever-decreasing significance. For if, as Hartman says, "we must free ethical thinking from 'common-sense' and give it wings to soar up to its own realm" and that "none has yet set out seriously to fashion the aerodynamic structures appropriate",<sup>7</sup> we must also remember the author Jack London's warning that philosophers "have left the real and solid earth and are up in the air with a word for a flying machine".<sup>8</sup> Our purpose, then, is to seek some answers to puzzling questions rather than to push curiosity to the limits for its own sake.

The branch of philosophy which is relevant to the understanding of evaluation is that called the Philosophy of Ethics. In fact the Philosophy of Ethics contains several branches, of which only one is of interest here, and it is as well to be aware of what these are.<sup>9</sup>

Prior to the Twentieth Century, the Philosophy of Ethics was primarily concerned with the advocacy and examination of moral theories embodying moral positions. This branch of the subject goes under the names of Classical Ethics, Moral Philosophy or Normative Ethics. More recently, philosophers have become interested in describing facts about people's moral views, known as Descriptive Ethics; and in conducting philosophical analyses of 'moral' terms and concepts, referred to as Philosophical Ethics or simply Ethics. It is to this latter sphere of interest that we shall turn.

The relevance of the Philosophy of Ethics (by which I generally intend to refer to the branch known as Ethics) to educational evaluation may be questioned. It is possible to establish relevance on four grounds.

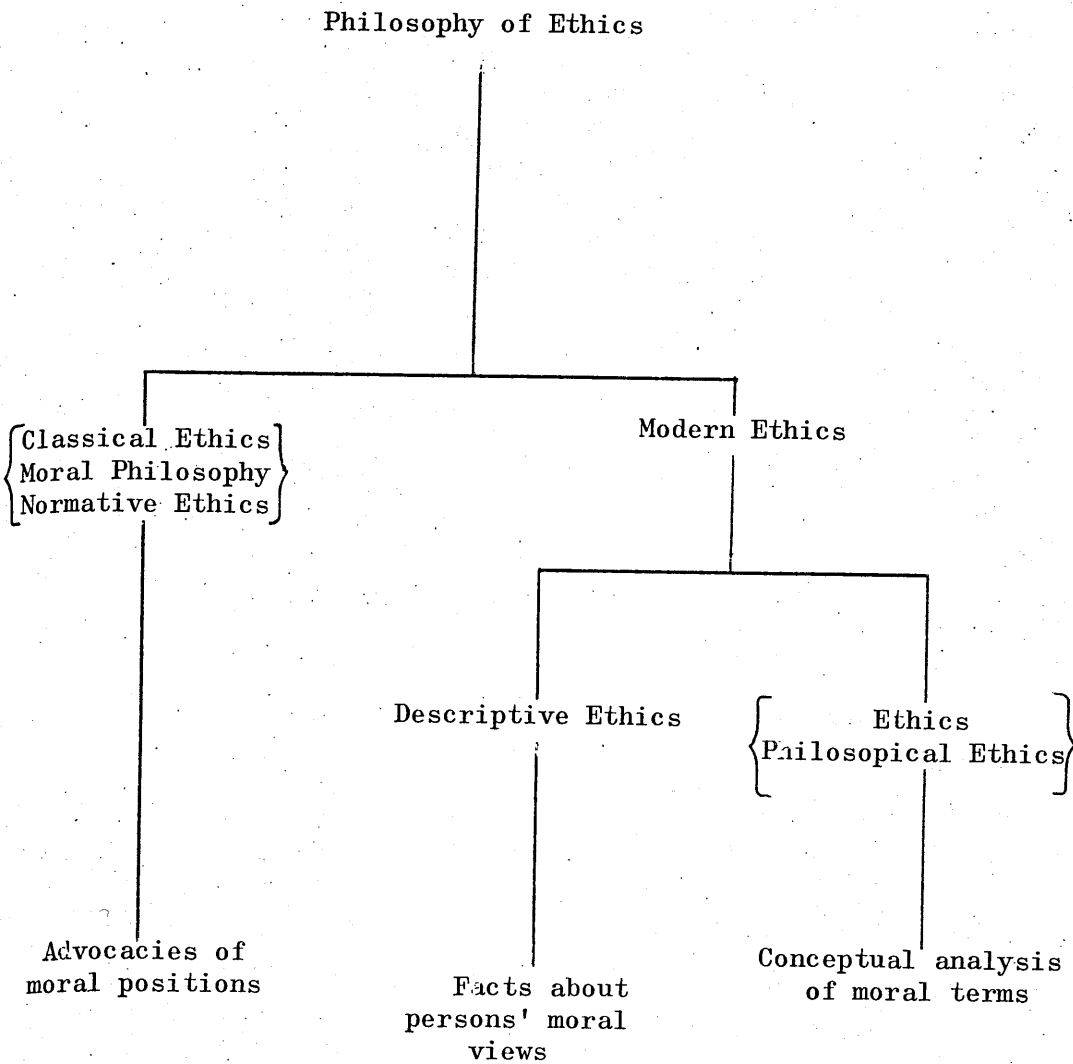
Firstly, educational evaluation is (or perhaps should be) concerned with questions such as 'Is this programme successful?', that is questions which concern value-judgement. Ethics is also concerned with questions of value and value-judgements, albeit at a general level.

Secondly, Ethics involves the clarification of ethical and value terms, and in doing this it is possible to derive clues as to how value questions can, or might be answered. For instance, a Utilitarian position implies that the proper way of determining something's value is to discover whether it gives people happiness. If this position, or some other position, could be firmly established, it would seem to give a definite direction to evaluative inquiries.

Thirdly, educational evaluation has often been viewed as an attempt to make evaluation 'scientific' and 'rational'. Ethics has also been much concerned with whether evaluative statements are of this kind, and thus is likely to offer some useful insights.

Finally, the current troubled state of educational evaluation, although arising from many causes, seems partly to arise from problems of conceptualisation. This suggests that an examination of the idea of evaluation is likely to be fruitful, and Ethics is a field which provides a body of thought on this subject.

Figure 2: Branches of the Philosophy of Ethics



There are, however, some difficulties in dealing with the contributions from the Philosophy of Ethics.

One of these arises from the historical development of the Philosophy of Ethics. As has been noted, the Philosophy of Ethics has traditionally concentrated upon the advocacy of moral positions, attempting to answer such questions as 'How ought Men to live?'. Philosophical Ethics, in contrast, has been concerned with the meaning of ethical terms. However, Philosophical Ethics has tended to use the classical views as the raw material for its own analyses, and has therefore focussed upon what is generally regarded as moral language, i.e. terms such as 'good', 'bad', 'right', 'wrong', 'duty', 'obligation', 'promises', etc. Educational evaluators, on the other hand, are more likely to be interested in concepts such as 'success' and 'effectiveness'. Certainly there is some lack of clarity in the philosophical literature about the range of applicability of conclusions about moral terms. It is hard to know, for instance, whether conclusions about 'good' also apply to words like 'success'. Fortunately, some recent work has considered evaluation at a general level which enables moral judgement to be seen as a particular 'special case' of evaluation. Urmson,<sup>10</sup> for instance, has worked from the mundane case of evaluating<sup>11</sup> trays of apples to arguments about 'good' and morality, and shows how these disparate issues may be considered from within the same framework. It is thus the general nature of evaluation, rather than moral evaluation in particular, which will be our main interest.

A second difficulty arises from the nature of the topic itself. Evaluation seems to be a particularly slippery customer, as anyone who has thought about it for a prolonged period will no doubt testify. Wittgenstein is reported as having said about the analysis of 'good' that it is "a terrible business - just terrible! You can at best stammer when you talk of it".<sup>12</sup> But for all this, it is hoped that it will be possible to say something clear and distinct about the problematic business of evaluation.

### Some Positions in Philosophical Ethics

Philosophical Ethics is concerned with the clarification of ethical or evaluative language, and views about the meaning of evaluative statements have implications for understanding how much statements might be constructed. The relevance of these views will be clearer when it is realised, as will be reiterated later, that the writer views evaluating as a process which leads to the formulation of evaluative statements, rather than, as some would have it, one which produces descriptions.

In analysing the meaning of ethical statements, philosophers have tended to ask what they are about, since one of the traditional ways of determining a word's meaning was to isolate its referents. Hence arguments have revolved around the question of whether such statements refer to their subject, to the speaker of the statement or perhaps to nothing at all. At the same time, philosophers have also been concerned with the verifiability of evaluative statements, partly because of what were thought to be the implications of the issue for ideas of 'justice', 'duty', and so on.

Philosophical positions on evaluation may be broadly divided into two families, those which assert that evaluative statements are verifiable declarative propositions and those which deny this. The former are sometimes called Descriptivist theories, and it is to these that we shall now turn.

Descriptivist theories may be either Naturalistic or non-Naturalistic. Naturalistic theories are those which assert that moral statements are statements about the world, and that they are equivalent in meaning to statements of non-moral fact. They are thus open to verification in the same way as 'scientific' statements are, since they are statements of fact. But what is it that evaluative terms 'describe'? One view is that they refer to the psychological state of the person uttering the evaluative statement. Thus 'x is good' is held to mean 'I feel good/My desires are satisfied', or 'Whenever I see x, I feel good/my desires are satisfied'. Evaluative statements thus refer to psychological states of the person, and are thus verifiable as other descriptions are.<sup>13</sup> An alternative view is that the referent of the evaluative term is the psychological state of a 'large number' of persons. On this view, 'x is good' means 'Most people feel good when they see x/ their desires are satisfied'. As before, this is held to be a verifiable description.

One of the implications of this view is that persons who are in dispute over the moral status of something are really arguing about nothing, since they are only saying 'I feel this' and 'I feel that'. These statements are not contradictory. Alternatively,



they are using argument to give an answer to an empirical question about how most people feel. Disputants thus think they are arguing about the characteristics of the phenomena they are judging, when in fact they are either just saying what their feelings are or giving their view on what most people's feelings are.

An objection to the first kind of naturalistic view is that if moral statements are equivalent in meaning to statements about feelings and desires, then 'Raping is bad, but it satisfies my desires' becomes 'Raping does not satisfy my desires, but it satisfies my desires', which leads to difficulties. If the second kind of view is accepted, then a minority of slaves kept by a majority of citizens would be incorrect to say that slavery was 'bad'. And this seems absurd.

In the writer's view, the resort to psychological states of persons as the referents of evaluative terms, constitutes a neat attempt at a solution to the problem of trying to specify the referents in terms of the phenomena being evaluated. An attempt to determine the referent of a term like 'black', for instance, could be approached by trying to isolate the common property of things said to be 'black', i.e. blackness. Yet the common empirical characteristics of all things said to be 'good' cannot be found. Thus to maintain the argument that 'good' stands for a property involves looking elsewhere, and thus to persons. The constant property then becomes either a state of a person, i.e. the common property of 'good' is a pleasant feeling, or the feelings of most people. And, of course, the problem of how it is that

people can disagree about the evaluative status of a phenomena when that phenomena presents exactly the same properties to each person, is solved by locating the difference in the situation in the person rather than the thing.

The Naturalistic position was largely abandoned by philosophers when Moore coined the Naturalistic Fallacy.<sup>14</sup> Moore asserted that attempts to define moral words with non-moral (descriptive) ones must always fail, since non-moral ones will always lack the essential element of moral meaning. This was, as it were, a reassertion of the unique character of evaluative language, although Moore himself was unable to go on to give a complete solution to the problem.<sup>15</sup>

There were, however, other positions which, in common with Naturalism, held that evaluative terms did refer to the properties of the world. Unlike Naturalism, the properties which were 'described' by moral words were non-natural. Thus one view holds that the world contains both moral entities, like 'goodness' and 'rightness', and natural entities, such as 'blackness' and 'horsiness'. Ethical statements have moral entities as their referents, but these entities cannot be observed as they are non-natural. Instead they are intuited. They are thus verifiable (in common with the Naturalistic view), but by means of intuition rather than observation. Thus Plato held that 'good' is there to be 'discovered' by men of knowledge, and that there is a definite, objective answer to the question 'What is truly good?'. A second Non-Naturalistic view holds that moral statements are 'expressions' of God's will. They are essentially descriptive, since 'x is good'

means 'God approves of x', which is a proposition about the psychological state of God! Although this cannot be verified by ordinary means, and perhaps not at all, it is still true or false since God exists and does or does not approve of things.

The Platonic view avoids the problem of locating the referents of moral terms in the natural properties of the phenomena. But instead of looking to the person, the notion of non-natural entities is introduced. In both the Platonic and theistic approaches, the ultimate source of authority about moral judgement rests with a specialised group of 'experts', namely 'men of knowledge' and ecclesiastics. The theistic position is subject to Moore's Naturalistic Fallacy, since it attempts to equate moral terms with non-moral statements about God, since Moore intended to rule out any equivalence with non-moral statements, regardless of whether they referred to the empirical or superempirical.

Neither of these approaches seem to be fruitful when related to real problems of evaluation. In the first case, 'men of knowledge' who are capable of evaluating intuitively are required yet do not seem likely to be found, and the metaphysical ring of 'moral entities' is conspicuously out of tune with the modern ethos. The abandonment of evaluation to ecclesiastics also seems an unlikely possibility.

The second family of theories is known as Non-Descriptivist. Whereas the Descriptivist theories hold that moral statements give information about the world or quasi-empirical phenomena, the Non-

Descriptive theories hold that evaluative terms have a different role than other adjectives, and that it is misleading to regard them as the names of properties.

The most prominent of these approaches is probably Emotivism. According to this view, ethical judgements are not descriptions of anything. Instead they are expressions of feelings, and are used to evoke similar feelings in others. As they are to be equated with squeals of delight, they cannot be verified. However, as Urmsen<sup>16</sup> has pointed out, whilst evaluative statements may be used as expressions of feelings, they cannot always be held to be such expressions, even if they use words with emotive meanings. Indeed, judges are often required to deliberately eliminate their feelings from their judgements. Therefore to regard all moral judgements as simply expressions of feeling seems to be inadequate.

None of these positions seem to give an adequate account of evaluation, even on their own terms as accounts of 'the meaning' of 'moral' statements. However, they do provide some useful pointers for the purpose of understanding the evaluative process. The Naturalistic view seems correct in emphasising a relationship between non-evaluative criteria and evaluative terms, but wrong in reducing these criteria entirely to subjective states. The Non-Naturalistic view seems to be correct in suggesting that evaluative statements are not simply equivalent to naturalistic descriptions, but unhelpful in replacing these with 'moral entities' and 'God's will'. Emotivism is correct

in seeing that evaluation may involve feelings, but incorrect in regarding ethical statements as always expressions of feelings. Before considering the relevance of these contributions to the writer's view, and some more recent positions, it is necessary to undertake our own analysis of the nature of the concept of evaluation.

### Preliminary Considerations of the Nature of Evaluation

It seems fairly non-controversial to say that evaluating is a human mental activity in the same way as describing or counting are. It also appears not to be identical with these activities, but is instead a process in its own right. The outcomes of this process is an evaluative statement, or a set of such statements, which may or may not be expressed by the evaluator. Evaluating therefore, can be seen as the process which leads to the formulation of evaluative statements.

A useful starting point in our examination of evaluation seems to be evaluative statements themselves. Given that it is possible to recognise some statements as evaluative (without which the investigation could not proceed), they can be scrutinised and compared with other non-evaluative statements with a view to identifying some of their distinguishing features. To do this, the notion of 'description' will be used as a kind of fulcrum against which evaluative statements can be levered. In particular, then, we shall be concerned with evaluation statements which seem to resemble declarative propositions.

One possible distinguishing feature of evaluative statements might be that they have a special syntax. Yet, as has been said above, evaluative statements seem to resemble declarative propositions. For instance, there appears to be no difference in the syntax of the statements,

'The apple is green'

'The apple is first-class'

Both appear to take the subject-predicate form, where the predicate appears to refer to a property of the subject. This resemblance does, of course, partly explain the search for naturalistic referents of evaluative terms. One might be tempted to say on the basis of this that the process of formulating a description is no different from the process of formulating an evaluative statement, and insofar as we may separate the doing of describing and evaluating from the meaning of descriptive and evaluative statements, there is an element of truth in this.

If evaluative statements do not have a special form, then they may perhaps always utilise special adjectives which are quite distinct from descriptive ones. Certainly some terms do seem to be used in an exclusively evaluative fashion, such as 'good' and 'first-rate'. But there are also terms which can be used both descriptively and evaluatively, as in,

'The diamond is hard'

'The examination is hard'

It does not seem, therefore, that special adjectives are a distinguishing feature of evaluative statements.

Another possibility is that evaluative statements may never be able to be said to be true or false, and we have already encountered Emotivism as a position which advocates this view. To tackle this possibility, we shall examine two ways in which a proposition may be said to be true or false.

A statement such as 'A brother is a male sibling', sometimes called analytic propositions, is true by definition. It is not a statement about the world, but is rather a specification of the equivalence of meaning between sets of terms. It is not verifiable by examining empirical cases of brothers and male siblings, but is true by virtue of conventional agreement on the identity of the referents of the terms 'brother' and 'male sibling'.

On the other hand, statements such as 'The man is black', sometimes called synthetic propositions, may be verified empirically. Given that the referents of the term 'man' and 'black' are different, then the statement is true if each term's referent occurs empirically and false if either or both do not.

To say that an evaluative statement cannot be said to be true or false is thus to say that it cannot be verified in either of these ways.

Since the Emotivist position is that evaluative statements are not 'sayings' but 'doings', they are presumably regarded as

neither true or false, just as a cry of joy can be neither true or false. As the writer does not accept this position as adequate for evaluative statements, the question of verifiability is still relevant.

If there were an analytically true definition of an evaluative term such as 'good', this would imply that there were a conventionally agreed equivalence of meaning between this term and some other terms. In that event, someone who said 'A green apple is a good apple' when someone else said 'A green apple is a bad apple' would be in the same position as someone who said 'A male sibling is a brother' when someone also said 'A male sibling is a sister'. In the latter case, the second speaker could be accused of speaking incorrectly; he would be wrong to equate 'male sibling' with 'sister', because we conventionally do not accept this definition. In the former case, we find nothing strange about the competing definitions. People who insist on defining 'male siblings' as 'sisters' are quite likely to find themselves referred for psychiatric treatment, or at least for a course in literacy; yet we find nothing strange about people holding contradictory definitions of terms such as 'good'. If the argument can be extended to other evaluative terms, it would seem that they cannot be held to have true analytic definitions.

If evaluative statements were synthetic propositions then their verifiability would depend on their being conventional agreement on the empirical criteria for the evaluative term. As has already been noted, even philosophers disagree about what these



might be, and it seems a mistake to assert that all evaluative disputes are really misguided arguments over one's feelings or the satisfaction of one's desires. In the absence of conventionally agreed criteria, it would thus appear that evaluative statements are not synthetic propositions which can be empirically verified.

The result of this is that evaluation seems to proceed in a chaos of subjectivity in which every evaluation is as good as anyone else's. However, the difficulty appears to be relatively easily disposed of once the notion of verification is examined.

One of the basic notions inherent to verification is that a proposition is either true or false and never true and false. Hence the troublesome problem arises when there are two opposite evaluations of the same phenomena and no way can be found of adjudicating between them. Either such statements must be placed outside the arena of verification, by regarding them as exclamations or expressions of feeling, for instance; or some adequate basis must be found for confirming one as true and the other as false. Since the statements cannot be both true and false, some overarching criteria have been sought after which are applicable to every case. The subjects of evaluation show such variation that such criteria cannot be found there; hence the turn to psychological states and God's will, since these are thought to be present in every case of evaluation.

It can be seen, however, that the possibility of confirming a proposition as true as dependent upon there being an agreement on the referents of the terms in the proposition. Thus to utter a true proposition is to represent the world according to the rules

of language. When I say that my brother is 'my brother', I speak truly because he is my male sibling and it is a legitimate use of language, according to our culture, for one's male sibling to be represented as 'brother'. If it were to be agreed that 'sister' meant male sibling, then the proposition 'He is my brother' would be false. To take another example, if the term 'solid' were held to mean 'impenetrable by any object directly perceptible to a human', then the proposition 'The table is solid' would be true. Yet since 'solid' is conventionally held to mean 'impenetrable by any object', and as particles can pass through the space between the table's atoms, the proposition is false. It is possible to assert the falsity of the proposition because most people would agree that 'solid' means that 'nothing will go through it'. We are able to legitimately maintain that 'The table is solid' is really false because the normal criteria of 'solid' do not specify objects directly perceptible to humans, and we can agree that particles fulfil the definition of 'object'.

If there were a community of language users who agreed to the more limited definition of 'solid', we would have to acknowledge that for them 'The table is solid' is true. This would not be to deny, nor to require them to deny, that particles can pass through tables. The truth of a proposition, then, is dependent partly upon the state of the world and partly upon the rules for the representation of the world in language. In this way, verification ceases to be an all or nothing matter, and becomes instead relative. The fact that propositions tend not to be experienced as true or false in a relative sense is a tribute to the effectiveness of the normative

regulation of language rather than to the 'absolute' character of truth itself.

In this way, the problem of finding overarching criteria of truth for evaluative statements can be circumvented. The view that there is, in many cases, an intimate connection between the characteristics of a phenomenon and the evaluative term which is attributed to it, is in the writer's view an entirely correct one. The fact that we cannot isolate the common properties of all the things said to be 'good' does not necessarily require us to look elsewhere for criteria enabling an evaluative statement to be declared true or false. Insofar as an evaluative statement is based upon a rule relating the characteristics of the phenomena to the evaluative term, it is open to verification, just as a description is.

There is an important difference, however, between the status of the rules for the use of descriptive and evaluative terms. Just as descriptive terms are open in principle to alternative definitions, though in practice only a few or even one definition is regarded as legitimate, so evaluative terms are open to alternative definitions and in practice are much less restricted. The corollary of this is that evaluative statements are open only to 'relative' rather than 'absolute' verification.

It should be made clear at this point that the writer regards many (but we cannot say all) evaluative statements as basically similar to descriptions in the manner of their

formulation (but not in their meaning). As Urmson<sup>17</sup> demonstrates in the case of apples, evaluative statements can be, and are, often formulated on the basis of rules which relate the evaluative terms to sets of descriptions. The difference which we wish to emphasise at the moment, though not always the only difference, is that the rules for the employment of evaluative terms are malleable in a way in which those for descriptions appear not to be.

Accepting that many evaluative statements are formulated on the basis of rules which, by embodying descriptions, leave them open to what we have called 'relative' verification, this fact reveals another significant property of evaluative statements. This is that the criteria for the employment of the evaluative terms in such statements are always open to questioning whereas those for descriptive terms are not. It would be strange to ask, for instance, why 'a natural elevation of the earth's surface rising to a summit' is the criterion for the employment of the term 'mountain'.<sup>18</sup> One would not know how to answer this other than by referring to convention. However, I would not be puzzled if I were asked why 'having four bedrooms' is my criterion for the employment of the term 'good' in respect of houses. This does not mean that it is always possible to answer such a question, no more than it would be easy for most people to verbalise the criterion for the term 'mountain'. But in the context of an evaluative statement; such questions are not eccentric, as they would be in the case of a description.<sup>19</sup>

It is therefore always sensible or intelligible to ask someone who makes an evaluative statement why his criteria are whatever they are. Indeed, the fact that we often do this reflects the fact that we recognise that such criteria are not fixed. In answering such a question we are justifying the criteria. Evaluative statements thus seem to have the property of being open to the justification of the criteria for the employment of the evaluative term.

The remarks made so far might appear to indicate that an essentially descriptivist position is being adopted in respect of evaluative statements. In terms of the way in which many evaluative statements are formulated, then there is certainly an analogy with the formulation of descriptions. To say that evaluative terms are descriptive predicates in the sense of having fixed referents is unacceptable for reasons already given. Yet the introduction of the notion of malleable criteria seems simply to leave evaluative terms as descriptive predicates of an unusual

and. However, the distinction needs to be made between the basis on which a statement is formulated (to be made clear for evaluative statements in the next chapter) and the meaning of the statement. An evaluative statement may be intended as an expression of feeling. These seem to present little difficulty in their formulation, and are the least interesting kinds of evaluative statement from our point of view. But it is not denied that this is what some evaluative statements are intended to be. In terms of meaning, it is accepted that not all evaluative

statements are to be equated with descriptions, nor in this case are they formulated in a manner similar to descriptions.

Otherwise, the writer does regard evaluative statements as similar to descriptions, with the exception that, unlike descriptions of the world, they are open to legitimate inferences about the speaker as well as the thing which is spoken of. A statement such as 'The apple is green' allows a listener only the inference that the apple has a particular property (excluding the exceptional case where the apple is not green, in which case something may be inferred about the speaker's perceptual abilities or his command of language). But a state such as 'The apple is good' is open to the interpretation that the apple has certain properties, and that the speaker likes or has pleasant feelings in respect of it. A particular evaluative statement may be interpreted in either or both ways, and a listener may always be uncertain as to which of these ways is intended. Thus evaluative statements seem to be open to inferences about the subject of the evaluation and the person uttering the statement.

Evaluative statements may therefore always be regarded as ambiguous and they are ambiguous in two ways. They are ambiguous firstly because they do not immediately signify themselves as referring to the speaker or the thing spoken of or both. They are ambiguous secondly because the criteria used to apply them to subjects are fluid. This means that it is always possible that the criteria which are assumed to be applied are not.

At this point we can summarise our conclusions as follows:

Descriptive Terms	Evaluative Terms
Have criteria for employment which tend to be prespecified	Have criteria for employment which tend to be open to the specification of the speaker
Have criteria for employment which tend not to require justification	Have criteria for employment which tend to be open to justification
Descriptive Statements	Evaluative Statements
Tend to be open to absolute verification	Tend to be open to relative verification
Tend to be limited to inferences about the subject of the description	Tend to be open to inferences about the subject of the evaluation <u>and</u> the speaker

Here we have couched the distinctions in tentative terms in order to avoid giving the impression of sharp divisions. It is apparent, for instance, that some terms which would normally be regarded as descriptive, such as 'large', 'tall', 'fat', do show an openness in their criteria for employment. However, we can go further in our investigation of evaluation by trying to see the reason for their being evaluative terms at all, and, perhaps, their necessity.

### The Need for Evaluation

Our contention so far is that an important class of evaluative statements are formulated in the manner of descriptions but are at the same time more than just descriptions. We noted also that they are likely to be formulated on the basis of variable criteria which descriptions in general do not seem to allow. It is necessary to ask, therefore, why such 'quasi-descriptions' should occur, and in what way evaluations are more than descriptions. It is possible to do this by considering the relation of experience to its representation in language.

A unique capacity of human beings is their ability to represent their experience symbolically by using words to stand for segments of that experience. Thus, for instance, it is possible to represent the experience of redness by the word 'red'. By virtue of the human biological constitution, it is not the case that all experiences are experienced as neutral. At its most basic, some perceptions are experienced as painful, as when one looks at the mid-day sun; whilst others are neutral, as when one looks at an orange; and still others are pleasurable as in the case of the orgasm. If humans were devoid of this capacity to experience even this primitively biological significance in respect of their perceptions, then a purely descriptive language would seem a perfectly adequate tool for representing the world. If everything were of equal significance then there would be no more to say that that things are as they were, if there were any point in saying anything at all. Given that not every state of



affairs is of equal significance, then it seems necessary to be able to have some way of expressing this very significance. Evaluative terms seem to do just this, by signifying not only a state of the world, but also the significance of this state. The need for evaluation arises, therefore, from the fact that men are capable of experience and are constrained by experience.

We can see that the possibility of sharing the meaning of a term such as 'red' is dependent upon our being perceptually homogeneous. Although we can never be sure that what we perceive as 'red' is the same as others perceptions (although we have every reason to believe it is the same), the term functions adequately for communication insofar as we can agree on whether something is 'red' or not. In effect we share the same criteria for the employment of the term 'red'. Where the significance of evaluative terms rests on the biological dimension of pleasure-pain, we might expect a similar consensus, since just as we are homogeneous in respect of visual perception so we are homogenous in our capacity to experience pleasure-pain. It therefore becomes hard to see how it is that different people evaluate the same phenomena in opposed ways. However, the pleasure-pain dimension is only the most primitive source for evaluation. A state of affairs may derive its significance from sources which are not so directly dependent upon our biology. Society's culture, the individual's social and psychological milieu, his role and social position, all give grounds for attaching a different significance to the same state of affairs, for different evaluations of the same phenomena.

The malleability of the criteria for the employment of evaluative terms is thus not an unfortunate accident arising from ignorance of what they 'truly' are; nor are the ensuing disagreements puzzling. The stability of the criteria for descriptive terms is based on our perceptual homogeneity; the variability of the criteria for evaluative terms is based on the heterogeneity of the significance which we attach to our perceptions. Some of the sources of this heterogeneity will be examined later, but it is obvious, for instance that a man with ten children will attach a different significance to a house with one bedroom than a man with no children. Thus we would expect them to evaluate houses differently.

We can say, then, that men do not only perceive and describe the world, but also attach significance to their perceptions most primitively by virtue of their biological capacity for experiencing pleasure and pain. Beyond this there are many sources of significance which are directed to experience. Evaluative terms serve to signify this significance, and derive their criteria for employment from the significance which the states of the world represented by these criteria have. Given that the significance of a state of the world varies between persons, this gives rise to variation in the criteria.

#### Summary of Philosophical Considerations

An examination has been made of some philosophical positions on the nature of evaluation, with a view to understanding what is involved in doing evaluation. The prima facie relevance of this

was stated in terms of the implications of the meaning of evaluative statements for the conduct of evaluation. Thus if evaluative statements were no more than a description of a person's feelings or the state of his desires, the proper procedure for dealing with a request to evaluate something would appear to be simply to consult one's feelings. Some difficulties in maintaining positions such as this were noted.

Evaluative statements were then scrutinised, together with some evaluative terms, and some interesting differences from descriptions were noted. In particular it has been suggested that evaluative terms tend to display an openness in their criteria for employment which is based upon the different significance which people attach to their perceptions.

There is a very great difficulty in all this in saying anything that is generally applicable to evaluation. We have already noted the difficulties which arise when any single view is used as a basis for making general pronouncements on evaluation.

But as our purpose is to provide a helpful foundation for doing evaluation, rather than, for example, determining whether or not someone is or is not uttering an evaluative statement, we intend to put forward a model suited to this end. Such a model does not purport to be a theory of how all evaluative statements are constructed, though undoubtedly many are constructed in this way. It does not therefore claim to encompass everything that we might regard as an instance of evaluation. It does however intend to be a representation of the way in which many important evaluations are

undertaken, and to provide a framework in which practical evaluation problems can be handled..

Before describing the model, it is necessary to state some of the assumptions on which it is based. These reflect the writer's stance in relation to the philosophical positions cited earlier. These assumptions are as follows.

Firstly, that although some evaluative statements may be no more than expressions of feelings caused (according to Stevenson)<sup>20</sup> by the speaker's emotions, many are not formulated on this basis.

Secondly, that many evaluative statements are formulated on the basis of information about the phenomena which is evaluated, and are thus dependent on the attributes of the phenomena.

Thirdly, that insofar as evaluations are dependent upon the attributes of a phenomena, they resemble descriptions in the manner of their formulation, with the exception that the criteria for employment tend to be open. They are thus verifiable in relation to these criteria.

Fourthly, that in the manner of their interpretation evaluative statements are unlike descriptions in that they allow legitimate inferences to be made about the speaker as well as what is spoken of. They may be interpreted as descriptions of the world and as indications of the significance the world has for the speaker.

A corollary of these assumptions is that 'objectivist' arguments about the referents of evaluative terms are rejected. Since the objectivity of descriptions depends on consensus about the criteria for the employment of descriptive terms, the establishment of the objectivity of evaluative terms is similarly dependent upon consensus. Attempts to erect such criteria have failed, either logically or pragmatically, where this has been held to involve creating criteria which are overarching. In other words, where 'objectivity' has been held to depend on universal agreement on the referents of a term, no such agreement has been forthcoming. Nor does it appear likely to come until everyone attaches the same significance to everything, a most unlikely possibility. This is not to say, however, that agreement on criteria never occurs, and indeed in these cases what is 'good' seems to take on the appearance of objective fact. Yet however, widespread the agreement, it will always be possible and legitimate for an individual to advocate criteria of a different kind.

Where evaluations are made on the basis of the attributes of the phenomena being evaluated, the writer believes that it is not possible to have a ready made specification of those criteria. Indeed, it is this very fact that helps to make formalised evaluations difficult to execute, and which presents the professional evaluator with his most formidable task. The criteria for 'good', 'average', 'inadequate', 'fair', and so on, are not ready made; they have to be constructed, and this, as will be shown, can be a very difficult matter. Even Hartman's<sup>21</sup> recent attempt to provide a calculus of value, in which a thing is 'good' if it has all the

properties given in the intension of its concept, does not provide an adequate practical tool, because it fails to recognise that the criteria for an evaluative term reflect the significance of a phenomena for an actor; and this significance cannot be overridden by an abstract definition relating value to a concept's intension. Evaluation would certainly be much simplified if criteria could be identified in this way. But as this approach only serves the purpose in certain special circumstances (as in the case of an exemplary instance of a concept), other methods of generating criteria must be found.

Having clarified our view on some of the philosophical arguments about evaluation, it is now possible to describe a model for evaluation and to examine some of its implications.

REFERENCES

1. For example, Stufflebeam's definition that "educational evaluation is the process of delineating, obtaining and providing useful information for judging decision alternatives". Stufflebeam, p.40.
2. R.H. Thouless, Straight and Crooked Thinking, London 1953, p. 89. My emphasis.
3. J.R. Sanders and D.J. Cunningham, 'A Structure for Formative Evaluation in Product Development', Review of Educational Research, 43 (1973), p.217.
4. Cf. Stufflebeam's 'Context Evaluation'.
5. See Scriven (1971) for criticisms of Stufflebeam's notions of evaluation. To be fair, we have already noted the latter's reservations about using the term. Similarly, Sanders and Cunningham do indicate that they have defined the term for their own purposes. We simply wish to point out its limitations as a general definition.
6. Cited by R.S. Hartman in The Structure of Value, Illinois 1967, p.75.
7. Ibid., p.87.
8. J. London, The Iron Heel, London 1947, p.18.

9. Works consulted on the Philosophy of Ethics include  
E.R. Emmett, Learning to Philosophize, Harmondsworth, Middlesex 1968; Hartman; J. Passmore, A Hundred Years of Philosophy, Harmondsworth, Middlesex 1968;  
R.H. Popkin and A. Stroll, Philosophy, London 1956;  
J.O. Urmson, ed., The Concise Encyclopaedia of Western Philosophy and Philosophers, London 1960; The Emotive Theory of Ethics, London 1968; J.O. Urmson, 'On Grading', Mind, 59 (1950).
10. Urmson (1950).
11. Urmson uses the term 'grading', but makes it clear that 'evaluating' might be used just as well.
12. Cited by C.H. Waddington in Science and Ethics, London 1942, p.7.
13. Some would argue that statements describing internal feelings are not verifiable, if this means observable by a multiplicity of observers.
14. "The work of Moore convinced most philosophers that naturalistic definition of moral terms could be ruled out". Urmson (1960), p.139.
15. See Hartman, Chapter Four.
16. Urmson (1968).
17. Urmson (1950).
18. The example is drawn from P.H. Martorella, Concept Learning: Designs for Instruction, London 1972, p.7.



19. This point has been made by both J.R. Coombs, 'Objectives of Value Analysis', in L.E. Metcalf, ed., Values Education: Rationale, Strategies and Procedures, Washington D.C. 1971, p.13, and J. Wilson, Thinking with Concepts, Cambridge 1963. Wilson writes; "Suppose someone said 'That's a good book' and we asked him 'What do you mean a good book?' This is a perfectly reasonable question: and it is also a question of concept, because what we want to know is what counts as 'a good book' with him". p.9.

20. See Urmson (1968).

21. See Hartman, p.101 ff.

## CHAPTER FIVE

### A MODEL OF EVALUATION

It is now possible to propose a model of the evaluative process. This model is of a general kind, and is intended to suggest an approach to evaluation which is applicable to a wide range of phenomena, rather than one designed for the evaluation of some specific topic, such as curricula. It therefore represents some structural components of evaluation which are given their substantive form in particular contexts.

#### Implicit and Explicit Evaluation

In our everyday lives evaluation is a common-place activity. Insofar as a person reasons at all over the choices which represent themselves to him, then he evaluates. As Urmson says, "grading and the application of grading labels are common activities ..... we all need to do it for the ordinary purposes of life."<sup>1</sup> Yet the approaches of everyday evaluation and the kind of specialised activity we can term 'professional evaluation' differ.

Professional evaluation is normally regarded as a formalised and explicit activity. Everyday evaluation is seen as informal and implicit. Indeed, one of the purposes of professional evaluation is the replacement of the everyday approach in certain significant contexts with a view to improving the quality of judgements and choices.

This distinction between the everyday and professional approaches is sometimes made in the literature on educational evaluation.

Wittrock, for example, says that "informal evaluations are judgements that do not necessarily involve an explicit statement of their bases, values, experiences, variables and data."<sup>2</sup> They occur, he says, when we judge the worth of our surroundings, our behaviour, changes in our behaviour, and events in our lives. The judgement is explicit but the bases are not. Formal evaluation, in contrast, involves explicit judgements and decisions and explicit statements and objective measures of the basis of the evaluation.

Suchman makes a similar distinction when he talks of evaluation as "a general process of judging the worthwhileness of some activity regardless of the method employed", and evaluative research as "the specific use of the scientific method for the purpose of making an evaluation."<sup>3</sup>

The importance of these distinctions for our purpose, is that in common with these authors, the writer is primarily concerned with 'professional' evaluations, normally those which are conducted by a specialist on behalf of a client. In these situations, it is suggested that an important condition to be met if the activity is to be carried through fully as an evaluation, is that the evaluative scheme (whose form is outlined in the model) be made explicit. It seems that the professional can only evaluate on behalf of the client if this is done. The model can thus be

seen as a programme or set of goals which must be reached in the conduct of professional evaluations. At the same time, however, the model can be seen as a representation of the underlying structure of many everyday evaluations, and hence as a framework for the investigation of their underpinnings.

The distinction between everyday and professional evaluation can be couched in terms of the use of implicit as opposed to explicit standards. The nature of standards will be discussed shortly, but for the moment we can say that they are rules which link information about an evaluative subject to evaluative terms.

In the case of evaluations based on explicit standards, the actor uses standards to enable him to formulate evaluative statements. Here the actor has in mind standards prior to being presented with an instance of the subject of evaluation, and this awareness has several important implications.

Firstly, he is able to direct his attention to particular aspects of the subject of evaluation, these aspects being specified by his standards. He thus knows what to look for when presented with a particular instance.

Secondly, he can make hypothetical evaluations, since an awareness of his standards enables him to express them as conditional statements, such as that "If x is the case, then y". He can thus project his evaluations beyond his current situation.

Thirdly, the actor who uses explicit standards is in a position to communicate them to others. This enables them to be made available, for discussion, and also enables them to be used by others. Hence the importance of explicit standards for professional evaluation, where the evaluator evaluates on behalf of a client. Similarly, when made explicit, standards can be used as specifications of goals and as specifications for design.

Where standards are implicit, the actor does not use his standards in a conscious manner, as the actor with explicit standards does. Such an approach might be characterised by saying that 'I know a good thing when I see it - but not before'. In this case evaluating becomes a personal accomplishment, and its bases are inaccessible. Where a professional evaluator is required to evaluate under such circumstances, then part of his task involves rendering the implicit standards explicit; only then can he use those standards on his clients behalf.

We are thus largely concerned with proposing a model for the conduct of explicit evaluations.

#### An Evaluation Model

As was noted earlier, evaluation can be seen as an activity in which the process of evaluating leads to the formulation of an evaluative statement. The current task is to show what this process entails.<sup>4</sup>

An evaluative statement can be seen to have both a subject, the phenomena which is being subjected to the evaluation, and an evaluative term. There are all sorts of phenomena which can be the subjects of evaluation, and it is proposed that evaluation proceeds on the basis of information about the subject. More specifically, it is proposed that evaluative terms are linked to descriptions of the subject by means of rules. We will call these rules evaluative standards. The evaluative standards give the criteria for the employment of evaluative terms, and these criteria are descriptions of states of the subject. When a set of evaluative standards has been formulated, the evaluator is equipped with a set of instructions which tell him which evaluative term to apply to any specific instance of the subject. To execute the evaluation he must have information about the specific instance in question. This description must then be located within the set of evaluative standards and the corresponding evaluative term applied. In the writer's terms, therefore, evaluating is the process of applying evaluative standards to information.

We envisage the items which constitute the subjects of the evaluation belonging to a particular class. The evaluative standards are set in relation to this class, and we thus get 'the evaluative standards for apples', '....motor cars', '.... doctors', and so on. Since the standards in question will refer only to the members of the class, it is necessary to provide a specification of the class such that items can be identified as legitimate or illegitimate candidates for evaluation under them.

We have said that the criteria specified within the standards are descriptions of states of the class members. This is not in fact always the case, as will be shown, but to assist clear exposition at this point we will assume that it is. Given this, the items within a class are envisaged as describable in terms of configurations of values of variables. These variables range over all the items in the class, and any specific item is describable as a subset of the combinations of values of the variables, where each value in the description refers to a different variable.

For the purpose of evaluation it is necessary to identify some variables (perhaps, in some cases, all of those which can be used to describe the items) as significant. It is the combinations of values of these variables which appear in the evaluative standards. When such variables are identified as significant, they can be called criterion variables. They are, as Meehan<sup>5</sup> puts it, normatively significant variables. That is, some variables may be regarded as significant as a basis for allocating evaluative terms, whilst others may be ignored. In the case of an economy, for instance, some people are vitally concerned about the 'level of unemployment', whilst others regard this as irrelevant in preference for the state of the 'balance of payments'. In one set of standards the variable 'unemployment' will appear but not 'balance of payments', and vice-versa. Formal construction of the combinations of the criterion variables generates a set of significant descriptions.

Once the set of significant descriptive categories has been constructed, it is necessary to impose an order on this set. In the case of a set of items such as apples, the order may be in terms of 'better-worse'. As a result of this ordering, we can note that the evaluator is in a position to evaluate any two items as 'better than', 'worse than', or 'equal to'. But the order of significant descriptions alone does not allow for the generation of an evaluative statement for a single item. To enable this, it is necessary to complete two further steps.

Firstly, a set of evaluative terms must be selected. As will be shown, there are usually some constraints on this selection, but the set will be something like 'good, fair, bad', 'excellent, good, average, fair, poor', 'first-class, second-class, third-class', and so on.

Secondly, it is necessary to define the rules which link the members of the set of significant descriptions to the evaluative terms. These rules constitute the evaluative standards.

The requirements here are that:

- a) Every member of the set of significant descriptions must appear in a rule linking it to an evaluative term. In mathematical terms, the evaluative standards must be a rule of correspondence for the set of significant descriptions and the set of evaluative terms. This means that for every case which may be located within the set of significant descriptions (which is constructed so as to cover any case we are likely to encounter), there is a corresponding evaluative term.



- b) No member of the set of significant descriptions may be linked to more than one member of the set of evaluative terms. In mathematical terms, the evaluative standards must constitute a function, whereby every member of the set of significant descriptions must be linked to an evaluative term, more than one member of this set may be linked to the same evaluative term, but no member of this set may be linked to more than one evaluative term. In other words, every case falling within a given descriptive category has one evaluative term corresponding to it, though cases falling within different descriptive categories may be linked to the same evaluative term. This avoids the situation where within a given set of standards two identical items, in terms of the descriptive categories, are attributed different evaluative terms.
- c) Finally, every member of the set of evaluative terms must appear in a rule linking it to the set of significant descriptions. That is, the evaluative standards must constitute an onto function. This means that no evaluative term is redundant in the sense that it is not linked to any descriptive category. This prevents there being evaluative terms which serve no purpose.

In summary, then, the evaluative standards must constitute an onto function of either a one-to-one or many-to-one-kind.

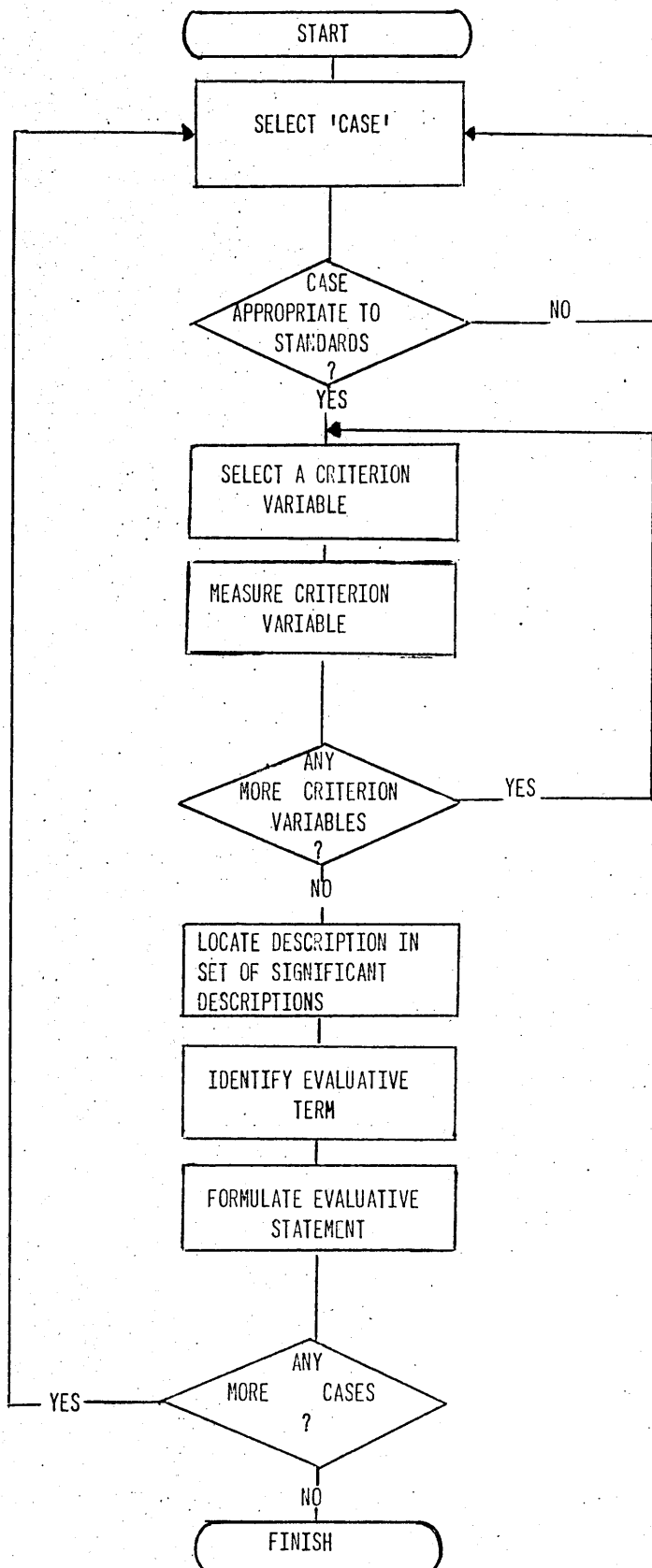
Once the evaluative standards have been constructed, one half, as it were, of the evaluative task has been completed. The evaluator is now equipped with a set of instructions which indicate how any actual instance of the phenomena which is being evaluated is to be

attributed an evaluative term. That is, he is equipped with a set of rules which enable him to determine its evaluative status. Up to now, however, no actual case has been evaluated. In essence a scheme has been created which has involved in its formulation numerous hypothetical evaluations, and the resultant standards in effect indicate how the evaluator intends to evaluate particular real cases.

In order to evaluate a specific case, the evaluator must locate the position of the case in terms of the values of the criterion variables which it takes, and then identify its position within the set of significant descriptions. This enables him to apply the appropriate evaluative term. This is the second 'half' of the evaluative task, the part which in certain circumstances requires research. Without standards, the information remains as simply a description; without information the standards remain as simply statements of intent. It is only by bringing the two together that it is possible to execute an evaluation, and for this reason evaluation can be regarded as involving two separable processes, namely standard-setting and information-getting. We will argue later that the predominant emphasis of current educational evaluation appears to be on the latter rather than the former.

We can now note some additional constraints imposed by the model, and some techniques for the construction of standards.

Figure 3: The Evaluation Process with an Established Set of Evaluative Standards



### Subject Variation

Firstly, we can notice that, in cases where items are evaluated on the basis of their characteristics, items must be able to be describable in different ways whilst remaining members of the class to which the standards refer.

Since evaluation involves discriminating among and ordering descriptions, the evaluative term set must contain at least two terms. If the only term available for evaluating a set of items were, say, 'good', then the possibility of evaluating disappears. It follows, therefore, that the set of descriptive categories which we have called the set of significant descriptions must contain at least two members. If this were not so it would not be possible to define a one-to-one or many-to-one onto function. This does not mean, however, that the actual cases which are evaluated must contain members which fall within each category. All the actual apples which are evaluated may turn out to be 'first-class'. But it is necessary to be able to conceive of a case which would not count as 'first-class'. This means that it must be possible to conceive of cases which manifest themselves as different values of a criterion variable. It follows, therefore, that in the sort of evaluative situation we are considering, if no way can be found of discriminating descriptively among items, so that they are from this point of view identical, such a set of items cannot be evaluated. Imagine, for instance, trying to set standards for molecules of hydrogen.

### Constitutive and Associative Variables

It is now time to qualify the statement about criterion variables being properties of the items evaluated. It certainly is the case that this is often so, but it is incorrect to insist that the criteria for evaluation necessarily refer to the item being evaluated. It is necessary to make a rather fragile distinction here. Let us imagine the case of an earthquake. Earthquakes are identifiable phenomena which show variation, for instance, in their strength, for which seismologists no doubt have a special term. It is possible to conceive of seismologists evaluating earthquakes on the basis of this property, and declaring some as 'first-class', or 'good', and others as 'second-rate', or 'bad'. And it might be that the criteria which they use are conceived of as properties of the earthquake. This is basically an illustrative example of evaluating a set of occurrences, (earthquakes) as we have considered it so far. Imagine also that an earthquake destroys a city, with huge loss of life. Someone might well evaluate such an earthquake as 'very bad', on the grounds that the earthquake resulted in the destruction of life and property. Yet these are not properties of the earthquake, but properties of a state of affairs which occurred as a result of the earthquake. The earthquake is 'bad', as it were, by association. It could be said, of course, that in this case the item being evaluated is not the earthquake at all, but its consequences. Yet we often do evaluate an item on this basis. Indeed, one theory of morals is that an act should be judged on the basis of its consequences rather

than in terms of the 'act itself'. Consider again, a ballet dancer who executes a particularly tricky manoeuvre. A critic may evaluate the movement as 'first-class' by virtue of his perceptions of the movement. The theatre manager, however, may judge it as 'poor' in that it did not produce thunderous applause from the audience. The non-applause cannot really be considered a property of the dancer's movements, yet it is quite intelligible for it to be evaluated on this basis. Thus we may distinguish between criterion variables which are constitutive of the items being evaluated, and those which are associated with the items. Thus criterion variables may be divided into constitutive criterion variables and associative criterion variables.

Urmson<sup>6</sup> makes a similar distinction when he talks of 'good of a kind' and 'good from a point of view'. In his example, he notes the case of a road built across a farmer's land. The road may be evaluated on the basis of properties inherent to roads; its camber, surface, drainage, materials, and so on. On this basis, it may be declared a 'good road' meaning good in respect of its properties. On the other hand, the farmer may judge it 'good' because it is positioned on an embankment which serves to shelter his previously exposed fields. In Urmson's terms, the road is 'good from the point of view of the farmer'. The farmer's criteria are not ones which refer to the properties of roads, but ones associated with this road by virtue of its consequences for him.

We must therefore qualify our earlier remarks concerning the source of criterion variables. They may not only be descriptive of the items to be evaluated; they may also be descriptive of phenomena associated with the items. Similarly, where an item is evaluated on this basis, it is not necessary to conceive of alternative descriptions of the items, only of the associated phenomena. Thus two identical earthquakes may in one case result in destruction of property (if it occurs in a populated area), and in the other not. Yet the two earthquakes may still be attributed a different evaluative status within the same set of standards. In both cases, however, viz. using constitutive or associative criterion variables, a specific item must meet the specification for class membership. It would be a mistake to declare an earthquake as 'bad' because it destroyed the city, if what destroyed the city was not an earthquake.

The value of this distinction is that it enables us to include in our model the sorts of situations which those theories which base themselves solely on the properties of phenomena being evaluated leave out. We can consider for a moment, Hartman's approach.

Hartman's useful contribution to the subject depends upon the notion of value as being determined by the extent to which a phenomena has the properties specified by its concept's intension. Unless we are to include the destruction which arises from an earthquake as part of the intension of the concept 'earthquake',

which would seem a doubtful thing to do, the destructive and non-destructive earthquakes must be valued equally on Hartman's terms. Since, as earthquakes, they have identical properties, they must have the same value. As we have seen, from one point of view this is satisfactory; seismologists might indeed formulate their standards in this way. But in the other situation, the descriptive categories are not directly referable to the item, and on Hartman's approach they would seem not to be relevant to evaluation.<sup>7</sup>

This distinction is of more importance in the case of the evaluation of courses of action, or means, so for the present we need only note that at least one variable must be selected as a criterion variable for the purposes of setting standards.

#### The Set of Significant Descriptions

The selection of criterion variables is obviously a crucial step in evaluation. For the present we will not consider how this selection is to be made, but once it is made, the set of significant descriptions can be constructed.

Since, in the case of a set of items, we have to account for every case which might arise, this can be done by constructing all the combinations of values of the criterion variables. In doing this all the logical possibilities, in terms of significant descriptions of items, are generated. Some of the resulting descriptive categories may turn out to be empirical non-starters, in which case they might be eliminated from the resultant set of significant descriptions. The set consists of all combinations of values of criterion variables,



where each element contains only one value of each variable at a time. That is, it is the cartesian product of the sets of values of the criterion variables. Thus with criterion variables 'colour: red, green', and 'shape: round, square', we generate the set 'green, round', 'green, square', 'red, round' and 'red, square'. Where an item can take more than one value of the same variable at the same time, as in the case of a two-colour item, the variable will appear twice in the set of combinations. It is the set of descriptive categories which is generated in this way which appears in the rules linking them to evaluative terms, and it is this set which requires ordering before the evaluative function can be defined.

A major problem that occurs here is that even quite small numbers of criterion variables, even with the minimal number of values, can generate very large sets of descriptive categories. If we take, for example, ten criterion variables each with two values, the total number of descriptive categories is  $2^{10} = 1024$ . A study of the criteria used to judge Open University counsellors, conducted by the author, produced fourteen binary variables.<sup>8</sup> This allows 16,384 combinations. When one considers that this set must be ordered in a non-random fashion, one begins to understand why evaluation is not necessarily an easy matter.

One way of looking at the problem of ordering the set is to regard it as involving the selection of one order from all the possible orders. Again the number of possible orders can turn out to be frighteningly large. In the case of 1024 descriptive categories, this is factorial 1024, namely  $1024 \times 1023 \times 1022 \times \dots \times 1$ .

A smaller example of two binary variables, gives  $2^2$  descriptive categories and  $4 \times 3 \times 2 \times 1 = 24$  possible orders. One of these orders must be specified as the order for the set of significant descriptions.

In fact the problem is even more complex than this, as the orders given above are only the possible total orders of the set. When a set is ordered in this way, every pair of elements is ordered; none of the elements are equivalent in the order. If we allow for equivalence, then the only requirement is that the set be divided into at least two parts, the parts constituting an ordered pair. Given this freedom, the number of possibilities increases substantially.

With a set of three descriptions,  $\{a, b, c\}$ , the number of total orders is as above, the factorial of three. That is six possible orders;  $(a,b,c)$ ,  $(a,c,b)$ ,  $(b,c,a)$ ,  $(c,b,a)$ ,  $(b,a,c)$ ,  $(c,a,b)$ . Allowing for equivalence, we must add a further twelve orders;  $(ab,c)$ ,  $(bc,a)$ , etc. Thus there are eighteen possible orders in all.

In order to calculate the number of possible orders which can be defined on a set of descriptions, we can use Pascal's Triangle, as shown overleaf. With a given number of descriptions,  $d$ , the procedure is as follows.

Figure 4: Pascal's Triangle and the Number of Ways of Ordering  
Given Numbers of Descriptions

Row Number	
2	.....1 1
3	.....1 2 1
4	.....1 3 3 1
5	.....1 4 6 4 1
6	.....1 5 10 10 5 1
7	.....1 6 15 20 15 6 1
8	.... 1 7 21 35 35 21 7 1
*	* * * * *

Number of Descriptions	Number of Orders
2	2
3	18
4	168
5	1800
6	22320
7	317520
8	5120640

1. Calculate the number of total orders = factorial  $d = a$
2. From Pascal's Triangle, locate the row with coordinate =  $d$ .
3. Sum the figures in the row and subtract 1 =  $b$ .
4. Multiply  $a$  by  $b$ .

As can be seen, the number of possible orders rises very rapidly with the addition of extra elements to the set of descriptions, starting with two possibilities for two descriptions, rising to 1800 with five, and more than five million with eight!

The requirement, then, is that the set of significant descriptions must be divided at least into two, and ordered. For the moment, it would appear that the number of criterion variables appearing in a given scheme needs to be kept as small as possible. There are some circumstances when this need not be so, as when each of the elements in the set of significant descriptions can be expressed as a value of some other variable which can itself be regarded as a criterion variable. Thus, for instance, if each of the values of the criterion variables can be expressed as a money-value, the combinations of money-values given for each descriptive category can be summed to give each element in the set a single value. All that remains is to order these money-values.

Approaching the problem of ordering the descriptive categories might best be done in the way suggested by Meehan.<sup>9</sup> Each criterion variable is taken separately, and its values ordered in respect of considerations which we have yet to make explicit.

Where only a single criterion variable is used, the ordering of this variable serves also to order the set of significant descriptions. In a multi-variable situation, each variable is ordered as if it were the only one appearing in the set of descriptions. Where a variable is inherently ordinal, the task may be simply to specify which 'way up' it is to run. Thus in the case of, say, 'temperature', order may be imposed in terms of 'the hotter the better', or 'the cooler the better'; this serves to order all the values of the variable. Where a variable is nominal, then each value must be located individually in the order.

We can notice here that it would be possible to go on to define the evaluative standards in respect of each criterion variable individually. In this case we arrive at a number of different sets of standards for the class of items, which allow evaluations to be made in respect of each variable. That is, we could evaluative the item as 'good, in respect of colour', 'poor, in respect of shape', and so on, without having to declare the evaluative status of the item as a whole. This certainly appears to be a viable solution to the problem of ordering complex descriptions, and avoids some of the difficulties of what Emmet calls the 'portmanteau value-judgement'.<sup>10</sup> It does not, of course, enable the formulation of a 'simple' evaluative statement.

In some cases, the ordering of the individual criterion variables can serve to impose a minimal degree of order on the set of significant descriptions. If, for instance, variable X with values  $x_1$  and  $x_2$  is given the order  $x_1, x_2$ , and variable Y similarly ordered  $y_1, y_2$ , then

it is sometimes possible to place the combination  $x_1, y_1$  first in the order and  $x_2, y_2$  last. This divides the set of descriptions into three parts with the order,

$$x_1, y_1$$
$$x_1, y_2 : x_2, y_1$$
$$x_2, y_2$$

The combinations in the middle are equivalent, which may or may not matter. But the ordering of the single variables has enabled the set of descriptions to be given a minimal degree of order. With more variables and more values, the same possibility holds. It is always likely to be easiest to decide what is to hold first and last places in the order, rather than the intermediate positions.

In other cases, the order of the individual variables does not enable the routine generation of even these best and worst worlds. Imagine, for instance, evaluating 'meals', where the criterion variables are 'food' and 'drink'. Assuming that the values of these variables are 'food: steak, cheese' and 'drink: whisky, wine', they might be ordered as 'cheese, steak' and 'whisky, wine'. The implied order of the descriptions is 'cheese, whisky' first, and 'steak, wine' last. But if consuming cheese and whisky make me ill, whereas steak and wine do not, the resultant order might be the reverse of that implied by the order of the individual variables.

Another approach to the problem of imposing order on the set of significant descriptions, each element of which consists of values of a number of different criterion variables, is to utilise numerical

indicators of the significance of both whole criterion variables and the values of individual criterion variables. In doing this, it is possible to construct a numerical label for each element of the set of significant descriptions, and since these numbers are members of an ordered set, the significant descriptions are ordered in correspondence with the order of this set.

There are several methods available for this approach,<sup>11</sup> but we will consider just one, devised by the author, which illustrates the general principles. The benefit of the approach is that it enables order to be imposed on the set of significant descriptions without the need for explicit comparison of pairs of elements, a process which is likely to be lengthy where the set is large. The difficulties arise over the assignation of numerical values, and over the assumption of additivity of these values. The approach would not, for instance, solve the problem encountered in the example of 'meals' given above. Nevertheless, the approach can be useful in giving a first approximation to order in the set of significant descriptions.

Such a procedure might operate as follows:

1. Impose a total or pre-order on the values of each criterion variable
2. Assign the number 1 to the first value of each criterion variable, and a number between 0 and 1 for every other value, such that if one value precedes another in the order, so does its numerical label.

The allocation of numbers must reflect the degree of significance which each value has. The first value will always be given the number 1, since it is the most significant value.

3. For the set of criterion variables, impose a total or pre-order, allowing that the whole set might be treated as equivalent. The elements to be ordered here are criterion variables themselves rather than their specific values.
4. Assign the number 1 to the first criterion variable and a number  $> 0$  and  $\leq 1$  to the other variables, so that if one criterion variable precedes another in the order, so does its numerical value.
5. Multiply the numerical label attached to each value of each criterion variable by the numerical label attached to the variable itself.
6. Construct the set of significant descriptions, consisting of the combinations of values of criterion variables.
7. Sum the numbers attached to the criterion variable values in each element, and order the set of significant descriptions in correspondence with the order of the numerical sums.

This approach embodies an attempt to measure the significance both of one criterion variable in relation to another, and of the values within each criterion variable. It thus attempts to reflect these structures in the numerical labels which are constructed for the elements of the set of significant descriptions, and hence in the order of the set.



It can be seen that:

1. If all criterion variables score 1, which is to say they are all of equal significance, the order will be dictated solely by the numbers given to the values of criterion variables.
2. If any criterion variable were to be given a score of 0, any significance attached to the values of that variable would have been eliminated via the process of multiplication. In effect this means that if a criterion variable is held to have no significance, it cannot be a criterion variable. Hence the restriction that every criterion variable must be given a significance  $> 0$ .

The difficulties attendant upon ordering the significant descriptions will be more or less great, depending on the particular circumstances that are relevant. For the moment all we can say is that the set must be given a minimal degree of order.

#### Evaluative Terms

It is worth considering here some of the types of evaluative terms, and to do this we summarise Urmson's<sup>12</sup> comments on the subject.

Urmson refers to evaluative terms as 'grading labels'. He notes that there are some labels which are used almost exclusively for evaluative purposes, such as 'good', and 'first-rate'. The use of these words enables a listener to recognise (with a few exceptions) that someone is evaluating. These he calls 'professional grading labels'. A feature of these labels, which

normally occur in sets, is that they almost always 'show their order'. That is, 'good' is regarded as higher than 'bad', and this order meaning is conveyed by the label itself. Thus the use of the term 'good' allows one at least to infer that whatever is given this status is not last in an ordered set.

Some professional labels have a general application. 'Good' and 'bad' are terms which can be applied to a wide range of phenomena. Others, however, are restricted to particular types of phenomena. For instance, 'Super Fancy' is a term which is used exclusively for grading, specifically, apples. Similarly, 'Hard Severe' is a term used for the evaluation of climbing routes by mountaineers. These restricted labels, says Urmson, tend to have more explicit criteria for their use than those having general application. We can also add that there seem to be some sets of evaluative terms which, while having a restricted range of application, do not have particularly explicit criteria. Terms such as 'wicked', 'saintly' and 'evil', for instance, seem to be restricted to humans or their actions; it would be most odd to call an apple 'wicked'.

A second group of evaluative terms are those which are used as evaluative terms in one context, but not in others. These are called 'amateur grading labels', and they are often ambiguous. 'Normal' is such a term, as in 'It is not normal to enjoy killing', which can be interpreted as 'Most people do not enjoy killing' or 'It is bad to enjoy killing'.

Finally, there are some words and symbols which are specially chosen as grading labels, and which do not have the obvious evaluative meaning of 'professional' labels. Nor do they have the ambiguity of 'amateur' labels. They can be virtually any word or symbol, such as 'A,B,C', 'red, white, green'. Their use does not immediately signify evaluation, and they do not necessarily have a natural order. They are called 'ad hoc' labels. One particular value of them is that they are emotionally neutral; as Urmson says, it is much less insulting to be told one is 'D+' than 'extremely poor'.

The remarks made in the next section about the calculation of the number of ways of specifying the evaluative function, assume that the order of the evaluative terms themselves is fixed. In some cases, this order is given naturally; it would be eccentric to have the order 'poor, excellent, fair, good, average'. Where this is not so, as with the last type of evaluative term mentioned above, ordering the terms is not a problem, since any order will do, provided that an order is fixed.

It can be seen that one of the useful properties of 'ad hoc' terms is that they usually allow an evaluative term set of any size to be constructed. Naturally occurring sets of evaluative terms seem to have only a few members, as with 'excellent, good, average, fair, poor', 'satisfactory, unsatisfactory', etc. They therefore restrict the number of evaluative distinctions which can be made. Their number can be increased by introducing qualifiers, such as 'very', 'fairly', 'somewhat', but their use can become unwieldy if a large number of evaluative terms are required. Thus it is

possible, but inconvenient, to use a term such as 'very, very, very, very good'. Since the evaluative term set must have order, it is convenient to use existing ordered sets, such as letters of the alphabet or integers, as 'ad hoc' evaluative terms.

The question inevitably arises as to whether it matters which evaluative term set is used in any set of evaluative standards. Does it matter, for example, whether the terms 'good', 'fair', 'poor' are used instead of 'A', 'B', 'C'? or 'Super Fancy', 'Extra Fancy' instead of 'first-rate', 'second-rate'? Could we not eliminate the process of selecting a set of evaluative terms by saying that all you need to do is select an appropriate number of symbols, each one different from all the rest?

Again it is necessary to refer to the context of the evaluation. It is obviously true that apple-graders could use evaluative terms other than 'Super Fancy', 'Extra Fancy', and so on. On the other hand, if you are evaluating boxes of apples, you face the possibility of being misunderstood if you use other terms and want to deal with other people in the fruit trade. If you were evaluating boxes of apples from your garden so that you could decide which ones to keep and which to throw away, the terms you chose to signify the significance of each box would not matter very much. However, if you had a large number of boxes, and had written a word on each one to indicate its significance, you would have to remember the relation of the words to the significance of each box. If the words had no 'pre-established' order, this would be inefficient and rather silly, when you could have used, say, letters of the alphabet.

Another constraint on the choice of the evaluative term set is the point of view from which the evaluation is made. The importance of points of view for evaluation is examined below, but we have already seen how evaluating from a 'moral' point of view suggests evaluative term sets including words such as 'wicked', 'immoral', 'promiscuous' and so on. Thus to evaluate from such a point of view but not use the evaluative terms such a point of view supplies, would be unnecessarily perverse.

We have also seen that evaluative terms of themselves are capable of having effects, independently of what they signify concerning the evaluation's subject. In so far as 'D+' and 'extremely poor' have the same criteria of employment for a given subject (such as a student's essay), then we might choose the set which includes 'D+' if we wanted to avoid hurting his feelings, and the set including 'extremely poor' if we wanted to administer an emotional shock.

For these reasons, which are not claimed to be exhaustive, the choice of evaluative term sets can be important, though what importance it has depends on the purpose of the evaluation. In the case of the student's essay, for example, if we don't care about the effect of the evaluative term on the student, it doesn't matter which set we use from that point of view. We may, of course, find other grounds for choice, such as that using the set including 'D+' means less writing than if the set including 'extremely poor' were used.

For these reasons, a consideration of the type of evaluative term set to be used is included as one of the activities in the model.

A final question, before we move on, is why is it necessary to introduce evaluative terms at all? Cannot evaluation proceed without evaluative terms? Certainly it seems it can, just as one can have feelings without having, or needing, words to express them. We are assuming, however, that the evaluator evaluates on behalf of the client, and that to do this in the fullest sense involves his being able to generate evaluative statements.

#### The Evaluative Function

Once order has been imposed and the evaluative terms selected, the final step in the construction of evaluative standards is the specification of the relation between the set of evaluative terms and the set of descriptive categories. As has been said, the set of descriptions must contain at least two elements, and the set of evaluative terms at least two terms. In this simple case there is only one way of linking the sets, as in fact there is in any case where the number of descriptions and the number of evaluative terms are equal. Note here that once the set of descriptions has been ordered, the number of elements effectively available for linkage to evaluative terms may have been reduced by grouping some of them together as equivalent within the order.

The set of evaluative terms may contain as many elements as there are descriptions available for linkage, but obviously not more. Once again there are many possible ways of specifying this linkage. The representation of Pascal's Triangle in grid form can be used to determine the number of possibilities. This may be done by locating the cell whose column coordinate is equal to the number of evaluative terms, and whose row coordinate is the number of descriptive categories available for linkage. The number in this cell is the number of possible ways of distributing the evaluative terms. As is clear, where the number of descriptions and evaluative terms are equal, there is always only one possibility.

Figure 5: Pascal's Triangle in Grid Form

Number of Descriptions	Number of Evaluative Terms							
	2	3	4	5	6	7	8	*
2	1							
3	2	1						
4	3	3	1					
5	4	6	4	1				
6	5	10	10	5	1			
7	6	15	20	15	6	1		
8	7	21	35	35	21	7	1	
*	*	*	*	*	*	*	*	

It is important to realise that although an evaluative function must be defined prior to the formulation of evaluative statements (since you do not know which evaluative term to apply to a given case without it), the content of this function may not be known until after the cases in question have been measured on the criterion variables.

It would be possible, for example, to specify that the top 10% of counsellors, when ordered by their number of years of experience of adult education, are 'first-class'. The number of years which would justify the application of the term 'first-class' then depends on the distribution of the population of counsellors on this variable. In this sense, the evaluative function is not completely specified until after data have been collected on the cases. However, it is still necessary to have a rule for generating the function in cases like this where it has not been defined independently.

Would it be possible, though, to specify an evaluative function without any reference to a criterion variable at all? Let us suppose there were 10 vacancies for counsellors and 100 applicants. We could tell the evaluator to designate the top 10% on any ordinal variable as 'suitable'. To do this, however, is to approach a random method of choice, in which the instruction would be to select 10 applicants at random. In both cases the requisite number of cases have been selected, but they have not been evaluated. The second approach, random selection, is arguably better than the first, since in the first, a measure is taken of



each applicant which is totally unnecessary, unless for one reason or another it is not 'respectable' to acknowledge that random selection is an appropriate strategy. And it is the case that a random strategy is justified in some circumstances.

Adopting such a strategy means that the 'choices' which result cannot themselves be justified. Indeed, if they could be, one's choice would not be made randomly. The choice between the strategies of choice themselves (random or reasoned) must, of course, be justified; the only justification for adopting a random strategy is because all attempts to find significant differences, or rather differences in the significances, have failed.

We mention that a random approach may be justified, because it seems, quite rightly, that we are in general reluctant to adopt this approach. A sensible strategy is always to search for some difference in the significance of alternatives, but if no difference can be found, it is better to choose randomly rather than by introducing an 'artificial' criterion variable. It is partly because of this possibility of creating 'artificial' evaluative schemes that we include the justification of standards among our meta-standards (see below). The fact that people feel embarrassed to have to admit that a 'choice' was made at random, does not mean that such a strategy is always inappropriate.

#### An Overview of the Model

We can now summarise the features of the model of evaluation.

Firstly, it is necessary to have a class of phenomena to be evaluated, the evaluation's subject. A specification must be given of the membership of the class, so that items for evaluation may be correctly selected.

Secondly, items must be capable of manifesting variation, or phenomena associated with them must be capable of showing variation. The selective representation of such variations is in terms of one or more criterion variables.

Thirdly, a set of significant descriptive categories is formed as the cartesian product of the criterion variables, or as the values of a single criterion variable.

Fourthly, this set is ordered either totally or partially to produced an ordered set of significant descriptions.

Fifthly, a set of evaluative terms is selected.

Sixthly, the evaluative standards (the function detailed above) are specified, linking the evaluative terms to the descriptive categories.

Seventhly, the specific items within the class for evaluation are described in terms of values of the criterion variables.

Lastly, the description is located within the evaluative standards, and the evaluative term identified. The evaluation is formulated as an evaluative statement which specifies the item's evaluative status.

It is probably worth running through an example here, to show how the model 'works' when applied to a particular case. In practice the order of the activities may vary from that given above.

The example to be used for illustrative purposes is a fairly simple one, and we will again leave the general question of how criterion variables are to be selected until later.

Let us suppose that we wish to set evaluative standards for motor-cars. 'Motor-cars' is the name of the class whose extension is all actual motor-cars. We firstly need to have a specification which enables us to identify motor cars from among all the other items which present themselves to us. To do this, we need to specify a minimal number of properties which conveniently define the class. This may well be a difficult goal to achieve, and must in any case depend upon the interpretative capacities of its user for its utility. Let us suppose that our specification for 'motor-cars' is 'a three or four wheeled object, propelled by motor, having seats for between one and five people, and primarily designed for the conveyance of people'.

Given this specification, we now need to have some idea of the variables which may be used to represent the members of this class. The specification given already provides three variables which range over the items - 'number of wheels', 'type of motor', 'number of seats' - and we could select any or all of these as criterion variables. But we may also describe the items in terms of 'colour',

'power to weight ratio', 'length' ("Will it fit your garage?"), 'country of manufacture', 'front or rear wheel drive', 'disc or drum brakes', 'drophead or fixed head', and so on. Clearly the range of variables applicable must be determined empirically, as well as the range of values of particular variables which the item may take. To do this, actual instances must be specified. There would be little point in considering 'gender' as a variable applicable to motor-cars; nor, for instance, considering lengths of one foot or fifty feet.

We will select two variables for designation as criterion variables, namely 'number of seats' and 'colour'. The number of values for 'seats' is already given as 'one to five'. Since we are interested in evaluating existing motor-cars, the range of values for the variable 'colour' can be determined empirically. Let us suppose they are 'red, white, blue'. The selection of these variables for inclusion in the evaluative standards means that all other variables will be treated as irrelevant for the purposes of evaluation. It can be noted that there is an intimate interplay between the empirical and the evaluative here. The fact that the number of seats a car has is significant, leads me to be interested in whether motor-cars can be described in such terms. At the same time, if I discover that motor-cars vary in their type of transmission arrangement, this may lead me to select this as a criterion variable. The search for information is informed by Value, and vice-versa.

By generating the cartesian product of these two variables, a set of significant descriptions is created. In this case, this set has fifteen members. Since the set has been created in a formal manner, it is possible that some of the descriptive categories have no empirical members. These elements can thus be eliminated. If, following this, any variable now presents only one value, the variables can be discounted from consideration; if all the cars are red, then colour ceases to be viable as a basis for distributing evaluative terms.

We will assume that only five elements of the set of significant descriptions remain: (one, red), (two, white), (three, blue), (four, white), (five, blue).

Taking the number of seats first, I decide that 'the more seats a car has, the better'. Thus the variable is ordered '5,4,3,2,1'. For 'colour', I choose the order 'red, white, blue'.

I now attempt to order the set of significant descriptions. Unfortunately, the element which comes first in terms of seats, comes last in terms of colour. Perhaps (one, red) and (five, blue) should be given an equal first position, followed by (four, white), with (two, white) and (three, blue) equal last. Or perhaps (four, white) should come first, with (five, blue) and (one, red) next, and (two- white) and (three, blue) last. Since there are 1800 possibilities, it would clearly be an impossibly difficult task to examine them all.

Unable to resolve this dilemma, the evaluator decides to attempt to give a numerical expression to the significance which he attaches to the values of the criterion variables. For the number of seats he attaches the following values:

Seats	Significance
5	1
4	0.9
3	0.5
2	0.2
1	0

His reasoning is that as he has a family of four, including himself, five seats would be best as he might occasionally want to take his widowed mother with him. As this is not likely to happen frequently, he could manage almost as well with four seats, as one child could sit on mother's knee. But with three seats, his wife would nearly always have to have one child on her knee, and with only two, one child would have to be left at home. With only one seat, he would always have to go out alone, and this would be intolerable.

As far as colour is concerned, he is not so particular, and he assigns the values:

Colour	Significance
red	1
White	0.8
blue	0.6

He also decides that there is nothing to choose between the criterion variables, and thus he attaches equal weight to them. Summing the numerical values for each alternative, he arrives at the following order:

Seats	Colour	Significance
4	white	1.7
5	blue	1.6
3	blue	1.1
2	white	1.0
1	red	0.9

The final task is the specification of the evaluative function. With five evaluative terms, there is no problem. With two evaluative terms there are four choices.

Exactly what the function will be, or whether indeed it is necessary to define a function depends on the purpose of the evaluation. The fact that we have not defined the purpose of the evaluation of the motor-cars is what lends the example its artificiality. We will consider in a moment the possible purposes of evaluation, and the implications of them for the technical problem of creating standards. Meanwhile we must emphasise that to ask someone to evaluate something without indicating the purpose of the evaluation and without equipping that person with an evaluative scheme, is to leave the way open for them to produce results which are of no relevance to the person making the request. Put another way, the request to 'evaluate (something)' contains very little information.

Let us now consider the case of Open University counsellors.

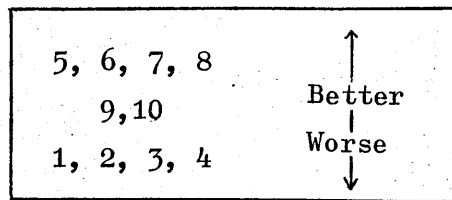
The class can be conveniently defined as 'all those persons contracted to work as Open University counsellors', the members being readily identified by reference to the University's administrative records.

By virtue of the fact that counsellors are human beings, a multitude of variables suggest themselves in terms of which individual counsellors could be described. Notice that we are talking about individual counsellors, not 'the counselling staff', and that the range of possible variables is different for each. For example, we might be able to consider counsellor in terms of variations in body shape (endomorph, ectomorph, mesomorph), but not the counselling staff. Similarly, we could describe the counselling staff in terms of its total membership, but not individual counsellors.

As in the case of motor-cars, it is necessary to select some variable or variables as significant. To do this, we need a context for the evaluation which will inform our selection. It may be, for example, that we want to be able to identify 'good' and 'bad' counsellors from the academic point of view, because we intend to replace the bad ones. The significant variable then becomes 'subject-matter-knowledge', or something similar.

Provided we can find, or create, a satisfactory indicator of 'subject-matter-knowledge', we then need to order this variable. Suppose the indicator is the result of a test scored from 0 to 10. We might order the scores as follows:





At first this may seem strange, since a score of 10 appears to be worse than a score of 5. However, we might reason that a score of 9 or 10 represents virtual infallibility and that this will be demoralising to students. Scores between 5 and 8 are enough to enable a counsellor to correct most errors without being able to correct all errors. Scores of 1 to 4 indicate that a counsellor would not be able to correct the majority of errors.

Given this order of the set of significant descriptions, we decide to use the evaluative terms 'good', 'satisfactory', 'poor', so that the evaluative standards are as in Figure 6. We can now proceed to acquire descriptions of actual counsellors, in terms of the criterion variable 'subject-matter knowledge', and determine each one's evaluative status 'from the academic point of view'.

#### Evaluating Courses of Action

Among the many phenomena that can be subjected to evaluation there is one class that is of special importance, namely courses of action. From the point of view of evaluation, the procedure for evaluating courses of action is basically that already outlined. Special problems arise, however, over determining the characteristics of the items in the evaluative subject set, which in this case are courses of action.

Figure 61. Hypothetical Set of Standards for Counsellors (from an academic point of view)

Subject
Open University Counsellors

If x (any OU counsellor) scores 5, 6, 7 or 8, he is a GOOD counsellor

Criterion Variable
Subject-matter Knowledge

If x (any OU counsellor) scores 9 or 10, he is a SATISFACTORY counsellor

From the  
academic  
point of  
view

If x (any OU counsellor) scores 1, 2, 3 or 4, he is a POOR counsellor

Indicator
Scores on subject matter test

We can assume that the search for a satisfactory course of action is usually initiated by the need or desire to attain some state of affairs which is currently non-existent. The implementation of a given course of action will involve the direct or indirect manipulation of those variables in terms of which the state of affairs is represented so as to convert their current values to those which are required. One of the important sources of criterion variables for a course of action is thus those which constitute the state of affairs which is to be manipulated. If, for example, we are evaluating the alternatives of withdrawing or not withdrawing all counselling services from the Open University, one of the criterion variables might be the likelihood of changing the rate at which students drop-out of courses, since this may be linked to the alternative course of action.

The evaluation of courses of action thus requires a model of the situation to which the course of action is relevant. The purpose of this model is to enable the projection of the outcomes of alternative courses of action. Each alternative is then described in terms of the values of the variables which constitute the state of affairs which it will produce, or is likely to produce, after implementation. It is these variables which will figure prominently among the criterion variables.

Insofar as the evaluation of courses of action may proceed largely in terms of the values of variables which describe a state of affairs resulting from their implementation, such evaluations can be undertaken in two modes.

In one situation we are faced, perhaps unexpectedly, with a choice of actions. We might be asked, for instance, whether we would like to go swimming or not. To evaluate the alternatives, we project the consequences of going as opposed to not going. If we go, we will get wet; if we don't, we won't. If we go, we will have to spend some money. If we don't, we won't. If we don't go, we might offend our friend who has asked us. If we do, we won't. And so on. The evaluative scheme would then be built around the variables which we identify as relevant to the alternative courses of action. The quality of the decision will depend on how accurate these projections are. In the same way, a decision on whether the Open University should or should not withdraw all counselling services could proceed in the same way, and is similarly dependent on reliable projections of consequences.

However, we may proceed differently. We may begin by constructing an evaluative scheme relevant to the state of affairs in question. Thus we might begin by defining a 'satisfactory' state of affairs without reference to courses of action. Having done this we may then search for a course of action which will achieve the state of affairs whose description falls highest in the ordered set.

We can apply the same idea to evaluating ordinary items. We decide the kind of item we want and then set about creating it. In this way, standards provide specifications which are relevant to design.

The evaluation of courses of action thus involves special problems which one would not normally associate with the evaluation of 'static' items. The descriptions which will be applied to courses of action are associated end states which follow their application. Where these relations are known and are deterministic, the problems are at a minimum. But in many situations relations may be only weakly established so that the consequences of pursuing a course of action are unpredictable.

The possibility of being able to associate courses of action with states of affairs is dependent on the one hand on the progress of the sciences in establishing relations between variables, and on the other on the skills of the operational researcher and cybarnetician. Both are concerned with the construction of models which enable projections of courses of action. Writers such as Kaufmann and Sargeant<sup>13</sup> provide examples of the use of 'linear programming' techniques which enable courses of action to be devised whose outcomes meet prespecified standards. The combinations of raw materials to give a specified output with specific characteristics is a typical case. Here the use of fairly simple graphical methods enables what is in effect a large set of significant descriptions to be divided into the 'satisfactory' and 'unsatisfactory' in a parsimonious manner.

Evaluators in education have not, it seems, been concerned with such problems on a large scale, perhaps because the well defined types of problems amenable to the operational research approach do not so frequently occur. Stufflebeam's approach is, perhaps, one

which most clearly recognises the problem of constructing and identifying courses of action as well as evaluating them.

### Sources of Standards

We have indicated that a central part of an evaluative scheme is a set of ordered criterion variables. We have also mentioned that there is a problem concerning the way in which these variables are selected and ordered. We noted earlier that in the case of a descriptive term, the criteria for employment are usually prespecified whereas in the case of an evaluative term they may not be. It seems that the criteria which are defined for evaluative terms cannot be arbitrary, since the very purpose of evaluation is to signify the significance of phenomena. The problem then is how to decide what is significant and what is not.

It is suggested that such decisions are made in respect of a number of general ordering rules. Some of these rules can be classified as Values, others as Preferences, although this may not exhaust the possibilities. Similarly, the notion of purpose in human affairs entails the notion of ordering states of affairs. If a person has a purpose he assigns priority to the state of affairs embodied by that purpose over other states of affairs. That is, having a purpose entails preferring the existence of the situation specified by the purpose over its non-existence.

Criterion variables may thus be selected and ordered in respect of these general ordering rules. Since they tend to be of a general nature, their relevance in the specific circumstances of an evaluation may not be immediately obvious. One of the problems

for the evaluator is to identify such rules as are applicable to the particular situation.

Doing this may involve the construction of fairly lengthy chains of reasoning, linking the subject of the evaluation to some principles of order. The way in which this is done is not likely to involve any simply described sequence of steps. Suppose, for instance, that I am asked to evaluate the set 'houses'. We will suppose that I am supplied with information about the ways in which houses can vary; the number of rooms, detached or semi-detached or terraced, number of storeys, styles of design, and so on. I will need some way of selecting from this set of variables some which I will designate as criterion variables.

The difficulty of considering examples is, of course, that they lack a real-world context. Similarly, open requests for evaluation lack the context which will inform the search for criterion variables. Any of these variables might be designated as criterion variables. It is only by having a real-world context that decisions can be made about which to select.

For instance, if my purpose is to have a house to live in myself, then the specification of better and worse houses will be informed by this purpose. I might begin by considering whether the number of rooms in the house are of any concern to me. I may simply have a preference for more rooms than less. I may be equipped with a cultural Value to the effect that 'It is bad to prefer less rooms to more'. Or I may have to make a lengthier search for

relevance. I may reason that if the house has more than a certain number of rooms, I will feel lonely. Given then a rule to prefer not feeling lonely to feeling lonely, this serves to establish the relevance and order of the variable 'rooms'. If I am unable to establish any grounds for the relevance of the variable, then it is excluded as a criterion. The search is informed therefore, both by the existence of ordering rules and purposes which the evaluator is aware of before he scrutinises possible criterion variables, and by information about the possibilities which stimulates the search for grounds of relevance. The construction of an evaluative scheme may thus start from either direction. Knowledge of the possibilities stimulates the search for some principle which can render them relevant; the possession of principles stimulates the search for information which will render them relevant. The fact that I care about the number of rooms I might be able to have, leads me to ask whether houses may be described in terms of such a variable, which, of course, they can. On the other hand, being told that houses vary in their degree of 'thermal efficiency' leads me to consider whether such a variable is significant to me or not.

#### Justification and Regress

Since Values, preferences and purposes are sources which inform the construction of an evaluative scheme, we envisage that the decomposition of such a scheme would lead to their identification. As we have noted, evaluative statements are always open to questioning. It is not eccentric to ask, for instance, why I designate a house



with ten rooms as 'good'. The answer to such a question can be termed a justification.

The first stage in obtaining such a justification would involve specifying the evaluative scheme. For instance, the answer be 'Because it has six rooms, and I prefer six room to any other number of rooms'. That is, that the criterion variable is 'number of rooms' and that the order is ' $6, \leq 6$ ' and that the function is ' $6, \text{good}; \leq 6, \text{not good}$ '. We can, however, go beyond this point and ask why the variable is relevant and why it is ordered as it is. The resulting justification can be pursued by further 'why' questions until a stopping point is reached beyond which no further justification can be offered. These stopping points are expected to take the form of statements of Value, preference or purpose.

Thus the answer to the question given above might be 'Because I prefer six-roomed houses to all other kinds', to which a subsequent question concerning why this preference is held may be 'Because I do'. Further questioning then becomes eccentric. Similarly, an answer such as 'Because everyone ought to want six-roomed houses' may admit no further justification. Similarly, we might justify a preference for existing as against not existing by a preference for the known to the unknown, but be unable to say why such a preference is held.

The link between evaluative standards and the underlying ordering principles which inform them is clearly important, since the value of the standards themselves is dependent upon this link. One of the ways of determining the adequacy of a set of standards

would be by scrutinising such justifications to see, in particular, whether the factual components of the justification are in fact valid. If, for instance, green apples are judged 'bad' on the grounds that apples with such a characteristic cause illness when eaten, then if this relation does not, in fact, hold, the standards lose their relevance. In this way, knowledge again plays its part in establishing adequate standards.

Considering the question of justification brings us to the problem of regress in evaluation. There is a saying about evaluation to the effect that:

Big criteria have little criteria upon their  
backs to bite 'em,  
The small ones have still smaller, and so on  
ad infinitum.<sup>14</sup>

The problem arises in that once a criterion variable is selected it is necessary to give its values an order. In effect, it is necessary to execute a sub-routine in which the problem of setting standards for the evaluative subject is replaced with the problem of setting standards for the values of the criterion variable, these values themselves becoming elements in a new evaluative subject set. Thus on the selection of the variable 'rooms' as a criterion in respect of houses, it is necessary to order this set. An ordering principle which is directly relevant to this set, solves the problem at once. Thus if the rule 'Prefer six rooms to all others' is held, the set can be ordered at once. On the other hand, it may be necessary to identify a criterion variable relevant to 'rooms', such as

'feeling of spaciousness' in terms of which each value of the variable 'rooms', now itself an evaluative subject set, can be described. Yet this variable now seems to become a new evaluative subject set; a principle for ordering this set must be found. The result of this is that a very complex structure of standards may underly the overt standards which are actually used to execute the evaluation of particular cases. Justifications for standards are therefore also likely to take on a similar complexity, constituting an unfolding of the submerged aspects of the scheme.

For example, suppose we were asking a client about the standards he would apply to counsellors. He might say that for a counsellor to be satisfactory, he must know the University's regulations. We could then ask why knowing the regulations is better than not knowing them, and pursuing each answer with further why-questions, something as shown in Figure 7 might follow.

However, it does not seem as though this regress must in practice be pursued endlessly. Rather we expect to find stopping points at which justification ceases. There is a point beyond which no further justification can be offered, and it is at this point the regress stops. From the point of view of scrutinising existing standards, as well as creating them, it is not, therefore, a necessarily hopeless task to pursue justifications. Indeed, as Harré and Secord<sup>15</sup> have pointed out in the context of problems of attitude change, it can be argued to be essential if a relevant strategy for altering attitudes is to be found. It also seems important for the conduct of explicit evaluations, as we shall see

Figure 7: Hypothetical Justification for a Standard for Counsellors

If a counsellor knows the regulations, he is satisfactory

If a counsellor does not know the regulations, he is unsatisfactory

If a counsellor knows the regulations, students will be advised  
correctly

If a counsellor does not know the regulations, students will not  
be advised correctly

If students are advised correctly, they will not make administrative  
errors

If students are not advised correctly, they will make administrative  
errors

If students do not make administrative errors, they will worry less

If students make administrative errors, they will worry more

If students worry less, they will study more effectively

If students worry more, they will study less effectively

If students study more effectively, they are more likely to get  
a degree

If students study less effectively, they are less likely to get  
a degree

If students are more likely to get a degree, the OU is more likely  
to continue

If students are less likely to get a degree, the OU is less likely  
to continue

Anything that helps the OU to continue is good

Anything that does not help the OU continue is bad.

when we come to consider standards which can be applied to standards themselves.

### Some Features of Values and Preferences

It is perhaps worth spending a few moments considering some differences and similarities among Values and preferences. To begin, we take an example of a statement of Value and a statement of preference. As with descriptive and evaluative statements, it is assumed that we can recognise cases of each.

For a statement of preference we have:

'I prefer eggs to cheese (Mozart to Beethoven, Daily  
Bugle to Morning Post, ....)'

And for a Value statement:

'Everyone should have a home to live in'

Other examples of Value statements might be:

'All men should be free'

'Everyone should have an equal chance in life'

'Parents should be respected'

'Men should subordinate their wills to the destiny of  
the State'

'Wrongdoers should be punished'

Firstly we can notice a similarity between Values and preferences in that both serve, as noted above, to resolve the problem of order. In the case of the preference, the rule serves to order the set containing eggs and cheese, and so on. Similarly, the Value 'Everyone should have homes to live in' serves to order the

set containing proportions of the population having and not having homes. In each case, the rules also serve as justifications for the evaluations which depend on these orders.

It can be seen that, in the case of Values, their very general nature creates a problem of establishing relevance. For example, the Value that 'Everyone should have a home to live in' says nothing about what is to count as a 'home'. Indeed, as Vickers<sup>16</sup> has pointed out, it is perhaps inherent in the notion of Values that their meaning remain open to the redefinition of succeeding generations. Preferences, by contrast, are more explicit, but at the same time more limited in their domain of relevance.

Preferences appear to be essentially personal, so that we accord a different status to them than to Values. Thus, we often preface a statement of preference with 'this is just my personal preference', by which I think it is implied that we do not expect any one else to agree with it. With a Value, we seem to be legislating for 'people in general'. Thus if I say that 'I prefer eggs to cheese', I am not saying anything about how I expect or want others to behave in relation to eggs and cheese. But if I say 'Parents should be respected', I seem to be saying that everyone should respect their parents. In this case I am claiming some degree of warranty for the ordering rule which I could not claim in the case of preferences.

It also seems that Values as rules are open to contravention whereas preferences are not. It is hard to see how one could 'contravene' a preference. One may mistakenly choose something one does not prefer, something lower in an order when something higher is available. Or one can change one's mind about a preference. But one cannot be held to have broken a preference rule in the way that one can be said to have violated the rule that one should respect one's parents. The source of legitimacy of a preference lies with oneself. There is an area in which one is free to prefer whatever one likes. But there is also an area in which the legitimacy of preference derives not from the individual, but from the culture in which he is enmeshed. The constraint of Values is external, whereas the constraint of preferences is internal. Hence the relative freedom to define preferences as one chooses.

The differential status of Values and preferences can again be illustrated by considering reactions to each when offered as a justification. If an act which is subject to a Value is in contravention of that Value, a justification by reference to preference is not likely to be accepted as legitimate. If I murder my wife, the justification that 'I prefer her dead to alive' is not likely to be regarded as legitimate in the face of the Value that 'People should not murder other people'. Values might thus be regarded as having prior claim over preferences as a legitimate source of standards where the subject in question can be brought under either with opposite results. If, however, action takes place in a context for which no Value applies, justification

by preference may be quite adequate. Thus if I choose to wear red ties rather than blue, it is perfectly legitimate to offer the justification that 'I prefer red ties to blue'.

Values and preferences are, then, important sources which inform the construction of evaluative standards.

### Points of View

The relevance of the notion of a point of view to evaluation has been noted by several authors, such as Coombs and Urmson.<sup>17</sup> When something is evaluated, it almost always involves the use of a point of view, and we sometimes acknowledge the relativity which this brings to the evaluation by saying that 'it all depends on which way you look at it'.<sup>18</sup> Clearly the question of whether the dropping of the atom bomb on Hiroshima was a 'success' will have a different answer depending on whether the evaluation is made from the point of view of the atomic scientists who designed it, or that of the inhabitants of the target area. We also noted earlier than an earthquake might be evaluated in relation to its strength, from the seismological point of view, or by reference to the loss of life it caused, which might be called the humanitarian point of view.

There are many other types of point of view; for instance, the moral aesthetic, medical, economic, administrative, vegetarian, educational, and so on. We would normally expect the incumbents of particular roles to be the advocates of specialised points of view, but they are nevertheless capable of being made publicly available. In effect, the kinds of coherent points of view



which we are able to identify by name provide a set of orientations to the task of setting standards. To evaluate from the religious point of view, for instance, in the case of say contraceptive techniques, closes off some variables as irrelevant, e.g. cost, and specifies others as criteria, e.g. having the sanction of Scripture. The medical point of view would produce a different set of standards, focussing perhaps on the ability of techniques to prevent contraception. Points of view may also supply the appropriate evaluative term set, as with terms such as 'immoral' and 'unprincipled' from the moral point of view; 'evil' and 'saintly' from the religious point of view; 'beautiful' and 'ugly' from the aesthetic point of view.

#### Levels of 'Sophistication' in Evaluation

Within the framework for evaluation that has been outlined, it is possible to discern various levels of 'sophistication' among evaluative schemes, reflecting increasing complexity. The more complex schemes are more difficult to construct than the simpler ones; but they also allow outcomes which the simpler ones do not. For example the simplest scheme only allows the judgement that one instance of a subject is 'better' or 'worse' than another in terms of a single characteristic. The most complex allows any instance a distinct evaluative status based on multiple criterion variables.

At the simplest level, we have the single criterion variable scheme without the evaluative function to the evaluative term set. As noted above, this arrangement allows the expression of the relation between any two items described by the criterion variable,

but not the determination of the evaluative status of a single item. The arrangement thus allows such statements as 'x is better than y in respect of p', but not that 'x is good (satisfactory, poor, etc.)'.

So, for example, the criterion variable 'cost per student' might be applied to the OU counselling service ordered in terms of 'the lower the better'. If we knew that the cost last year was £x and that it is now £2x, then the counselling system this year is worse than it was last year. Given this evaluative scheme, this is all we are able to say.

We can use this type of scheme with as many different criterion variables as we like. We might have added 'average number of complaints received by Senior Counsellors about counsellors' and 'ratio of male counsellors to female counsellors' to 'cost per student', giving three criterion variables ordered in terms of 'the lower the better'. This scheme allows three judgements of the counselling system this year as compared with last year. It may be better this year in terms of cost; worse in terms of sex ratio (and thus liable to attack from feminists); and no better nor worse than last year in terms of complaints. The comparison need not be with the same system at different points of time. Comparison could be made with, say, the University of Ruritania's counselling system, provided that it meets the criteria of membership of the class to which the OU counselling system belongs.

At the next level, there are multiple criterion variables formed into a set of significant descriptions representing their combinations, which is ordered but without the evaluative function.

This allows statements such as 'x is better than y', but not the determination of the evaluative status of single items.

Suppose that we had identified two criterion variables for the counselling service, such as total (money) cost per annum and total number of student complaints about the counselling service. Both of these might be ordered in terms of the lower the better. If these are the only significant variables, the Cartesian product is the universe of significant descriptions, which might look like this after ordering.

Total cost (£M)	Total complaints	
1	100-100	better ↑ ↓ worse
1	200-299	
2	100-199	
2	200-299	
1	300-399	
2	300-399	

This allows any given description to be declared as better/worse than any other (excluding itself), which is the same situation as with the single criterion variable. This scheme is more 'sophisticated', however, because it requires more than just the ordering of single variables, as in the previous two cases. Rather combinations of the values of the variables must be ordered which is a much more difficult process, depending on the number of variables involved.

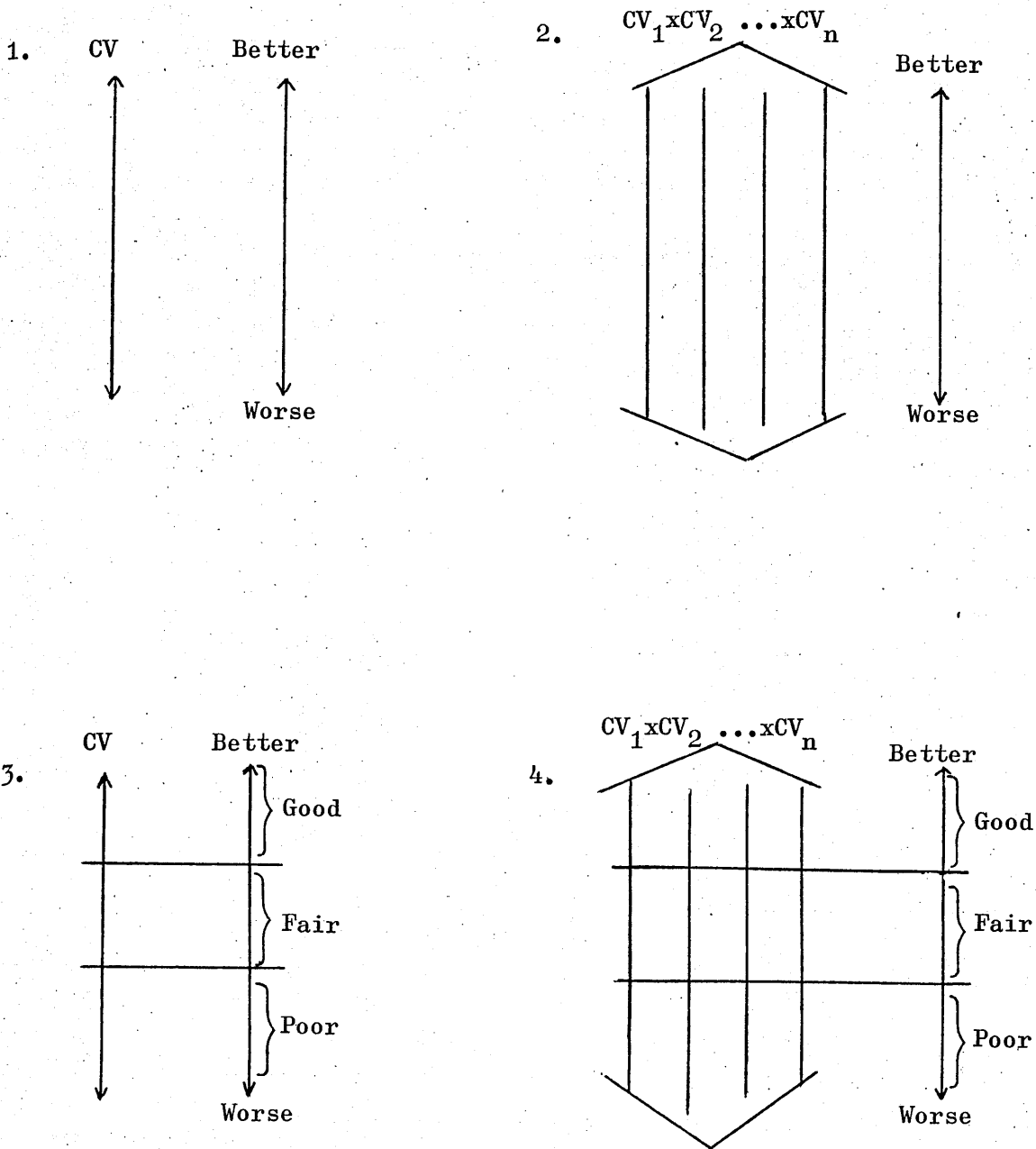
The next arrangement is the same as the first but with the evaluative function defined. This is really the simplest type of fully fledged evaluation, as we have defined it. Here, an evaluative statement can be formulated for any item. So we might have:

Cost per student (£'s)	
< 7	Very Satisfactory
7-8	Fairly Satisfactory
9-10	Fairly Unsatisfactory
> 11	Very Unsatisfactory

This scheme requires only the ordering of one criterion variable, but now it is necessary to make divisions within this order so that values, or sets of values, are linked to evaluative terms. It is no longer enough to specify that a value is better or worse than another; we now have to say at what point cost per student, as a measure of the counselling service, ceases to be 'very satisfactory' and becomes 'fairly satisfactory'.

The most complex evaluative scheme is similar to the second but with an evaluative function defined on the set of significant descriptions, these being combinations of the values of the criterion variables. This enables any item to be evaluated as 'good', 'poor', and so on. It allows the most comprehensive judgement of the subject, and at the same time the most considered judgement. If it were possible to include in such a scheme all the conceivable criterion variables for a subject, then the set of significant

**Figure 8: Levels of 'Sophistication' of Evaluative Schemes**



CV = Criterion Variable

descriptions would contain every significant description of the subject. Given that we have ordered this set and created an evaluative function, we will have taken the maximum amount of information about the subject into account in forming our judgement.

Which of these schemes is practical and desirable depends upon the context of the evaluation. The most comprehensive scheme is the last, but it is also the most difficult to construct. However, the more complex schemes often embody the simpler ones, in the sense that you have to construct the simpler scheme en route to the complex. For example, to construct scheme three, you have to construct scheme one, and to construct scheme four you must have constructed scheme two. The question is, when is it necessary to construct the various levels of scheme? Might it be, perhaps, that we rarely need to construct the most sophisticated scheme? And is there any reason to believe that one sort of scheme is preferable to another in some circumstances? There seems to be little point in creating the more complex schemes if simpler ones serve the purpose just as well.

This brings us back to the question of the purpose of an evaluation. Earlier we spoke of purposes informing evaluative schemes. If my purpose is to cut down a tree, this will influence the standards I apply to the axes available to me. If I intend to break down a door (as a fireman might), different considerations would apply, and the evaluative standards for the axes would change.

Purposes in these cases serve to affect the content of the evaluative standards. However, we are now referring to the relation

between the purpose of the evaluation and the structure of the evaluative standards, in terms of a scheme's degree of sophistication. In what sorts of situation, for instance, do we need to be able to declare something as 'good', rather than simply as 'better than' something else?

An important feature of the first two schemes is that no evaluative function is defined. The kind of statement which would result from the application of these sorts of scheme, is that 'x is better than y', but not that x is necessarily 'good'. The scheme would enable me to order the cases of the subject but, that is all.

Such schemes might be used when we wanted to know simply whether something had improved. We could establish this without having to say whether the existing state of affairs is 'satisfactory'. We might also use this sort of scheme when our sole purpose is to establish an order among cases, an order which is significant. For example, if we wanted to award prizes to contestants, we need to be able to place them in an order and match this with the order of magnitude of the prizes. We do not necessarily need to declare any particular positions in this order as 'good', 'poor', or whatever.

When might we need to introduce more than one criterion variable into such schemes? If we wanted to see if a system, such as the Open University counselling system, had improved over the years, we could use several schemes at level one. So we might find that it had improved with regard to the provision of vocational guidance, but had not improved in preventing students from dropping

out of the University. A number of schemes at level one would therefore be adequate if we wanted to discover which aspects of the system were moving in which directions.

A scheme at level two would be needed only if it were necessary to say that the counselling system 'as a whole' had improved.

Schemes at levels three and four differ from the less sophisticated ones mainly in that they involve associating particular values of criterion variables with evaluative terms. How we establish this relation depends on circumstances. Where we are evaluating a number of cases, the characteristics of the set may determine the characteristics specified by the function; or we can specify the function without reference to the characteristics of the cases. It seems that our reasons for wanting to specify evaluative functions must be because the evaluative subject can take on certain critical values. There must be some specific value which has particular significance. For example, the higher my income, the better, but only an income greater than £x p.a. is satisfactory, because with less than this I go hungry. Thus if my income is £x p.a. or greater, this is a satisfactory income, and if it is less than £x p.a. it is an unsatisfactory income.

However we arrive at these critical values, their existence suggests that we intend to behave differently towards cases falling into the different bands which they demarcate. If we have the power to change the counselling system at the Open University, it is not



enough to know that the system has improved since last year.

It is necessary to know whether the system has changed from being unsatisfactory to satisfactory. All the significant descriptions which fall below the critical value in the ordered set are unsatisfactory, and we seem to be saying that no matter which one of those descriptions represents the system as it now exists, we should look for ways to make a change.

It therefore seems that a judgement as to whether or not something has improved can be independent of a judgement of its evaluative status, but in deciding whether to improve something, we often need to have a way of identifying this. The occasions on which we can decide to devote resources to improving something regardless of whether it is very satisfactory, or very unsatisfactory, are few.

As with the simpler schemes, it is possible to devise a number of schemes at level three to apply to the same subject. These would serve for deciding whether or not to try to improve various aspects of the system.

The fourth and most complex type of scheme is required when an overall judgement is wanted, and more than one criterion variable must be taken into account. If we found that the Open University counselling service was 'effective' in giving vocational guidance, 'ineffective' in preventing students from dropping out, and 'fairly effective' in identifying students' academic problems, we could still be asked whether it is an effective system in general. It is hard to see what point there would be in asking such a question, but if we

wanted an evaluator to give the answer, it would be necessary to create a scheme at level four.

### Standards for Standards

Given that professional evaluation requires fairly explicit standards, it is possible to suggest some standards which might be applied to standards themselves. To do this is to suggest that some standards can be envisaged as more or less adequate than others. This entails the notion that some evaluations are better than others, their quality being dependent upon the quality of the standards.

From the point of view of the professional evaluator, the minimum condition which standards must meet is that they be explicit. This means that the significant descriptive categories must be so defined as to be empirically useful, and that the rules linking them to the evaluative terms must be specified. The evaluator need not go beyond this. However absurd the standards may appear to be, the evaluator is nevertheless equipped to conduct proxy evaluations.

From the point of view of the clients, standards would seem to need to meet other criteria of adequacy, such as that they should be justifiable and relevant. Whether the evaluator should be concerned with helping the client to realise these standards is a matter for argument. Stake, for instance, says that "whether people should think more rationally is not the issue since it is not the evaluator's task to reform human-judgment processes."<sup>19</sup> Yet it seems that the application of poorly formed standards to high quality data is just as bad as the application of well formed

standards to poor quality data. Neither results in a particularly happy state of affairs.<sup>20</sup>

One of the traditional aims of the formalisation of such processes as decision-making and evaluation has been to make them more rational. There are various views as to what might be regarded as 'rational',<sup>21</sup> but it is accepted, for instance, that decisions are likely to be more rational if the decision-maker is equipped with valid information about the alternatives. The task of the applied sciences, operational research, and so on, can thus be seen as to help to make decision-making a more rational process.

In the same way, writers such as Coombs regard value judgements as open to the application of standards of rationality. Thus, in that particular case, one of the aims of teaching students about the process of value judgement is to help them be more rational in their evaluations. This presupposes, of course, that it is better to make more rational value judgements rather than less rational ones. Thus Coombs gives the following standards for rational value judgement together with their justifications.

1. The purported facts supporting the judgement must be true or well confirmed. Since value judgements are partly based on factual considerations, the truth of the facts is of obvious relevance to the viability of the judgement.
2. The facts must be genuinely relevant to the person making the judgement. A person misrepresents his judgement if he does not really consider the variables involved as being of any significance. It is not rational to allow one's judgement to be influenced

by facts that one really believes to be of no importance.

3. The greater the range of relevant facts taken into account in making a judgement, the more adequate the judgement is likely to be. Persons who are unaware that the subject of the evaluation does present itself in ways which are significant to them, cannot take these facts into account. They may thus make judgements they would not have made had these facts been available.
4. The value principle implied by the value judgement must be acceptable to the person making the judgement. One cannot accept a value judgement and reject the principle implied by it without logical contradiction.

Whilst these standards are intended to apply to evaluative statements, and certainly appear reasonable from that point of view, we may also suggest standards which are applicable to evaluative schemes as a whole. Thus it can be suggested that evaluative schemes should be:

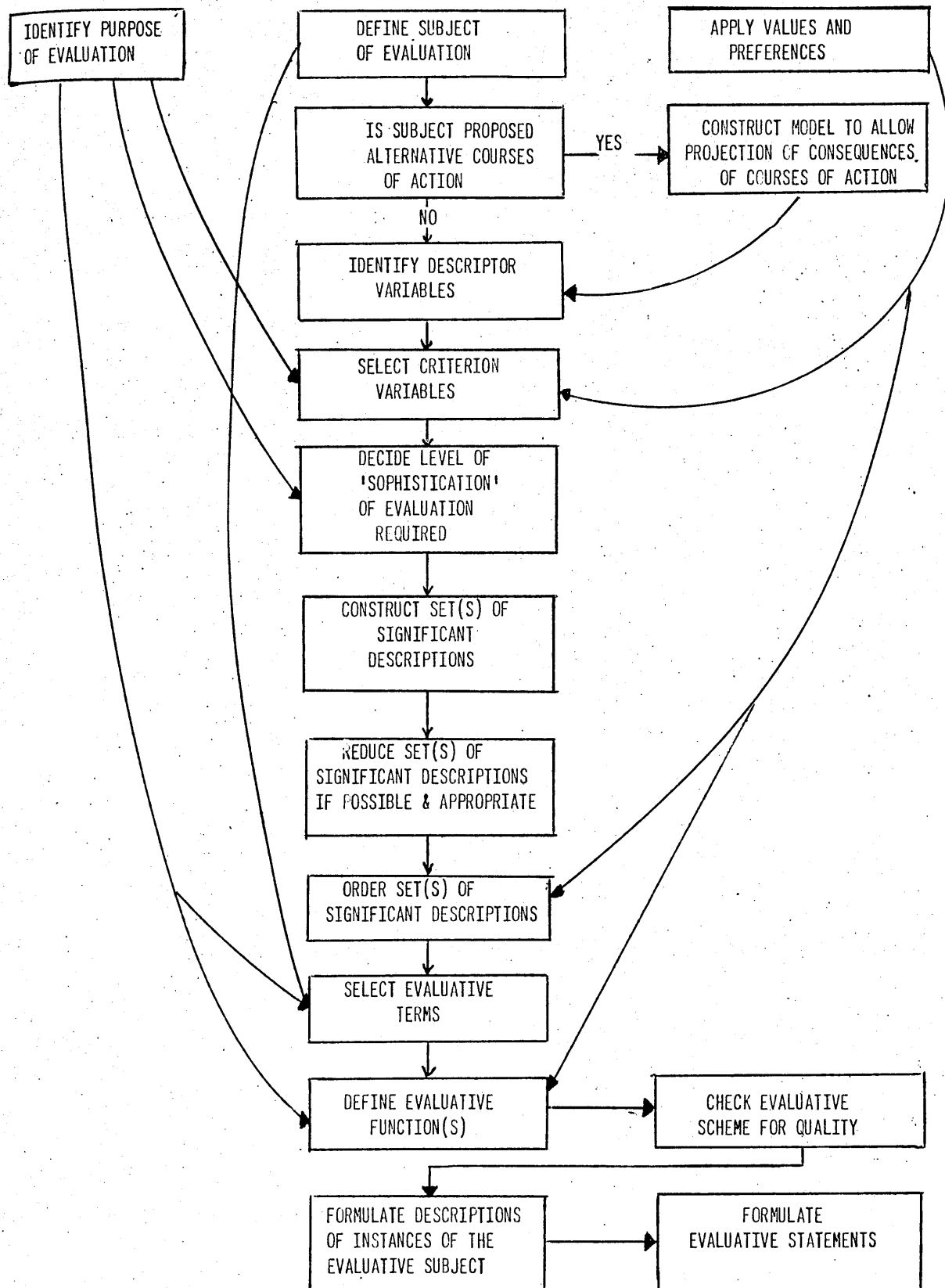
1. Justified rather than unjustified, and that the justification should be logically and empirically valid. Justified standards are essentially reasoned standards. Unjustified standards are unreasoned and are basically arbitrary. It is thus hard to see how such standards could be regarded as adequate. At the extreme, they would be created in a random fashion. They would not be informed by Values, preferences nor purposes, and would constitute an arbitrary device for formulating evaluative statements.

Most standards are capable of some sort of justification. These justifications can be expected to take the form of chains of argument consisting of both factual and evaluative assertions (see below, Chapter Six). They can be tested on logical grounds by examining the reasoning which links the justifications to the standards. Similarly, they may be tested on empirical grounds by examining the validity of any factual claims that are made. Among educational evaluators, Sanders and Cunningham<sup>23</sup> have proposed similar tests for application to instructional objectives.

These criteria of justification are purely technical. If they are met, this does not necessarily mean that they are adequate justifications on other grounds. They may still be disputed, but externally rather than internally.

2. Comprehensive rather than partial. Standards should be informed by all the Values, preferences and purposes which can be established as relevant. This is analogous to Coombs' criteria 2 and 3 above. If a set of standards does not reflect all the ordering principles which are relevant, any information about the evaluative subject which is relevant to these principles is ignored. Obviously if the 'colour' of motor-cars is significant to someone wanting to choose a car, he should include it in his standards. If he does not and there are two cars available at the top of his list which vary only in colour, the colour of the car he actually gets will be decided at random - that is if an evaluator is evaluating on his behalf. In the case of professional

Figure 9: Summary of Tasks for Explicit Evaluations



evaluation, it is important to identify as many possible criterion variables as can be before conducting a major data-collection exercise. Otherwise, the data provided may fail to inform some crucial criterion variable.

3. Specific rather than vague. Since the aim of the formal evaluation exercise is to determine the evaluative status of an empirical phenomenon, it is important that the descriptive categories to be applied to it should be specific and unambiguous. Without this, there will be uncertainty as to whether an instance is to be described in one way or another, and hence doubts about its evaluative status.

It is not to be expected that these standards for standards can be easily met. Very much depends upon the context of the evaluation, what is being evaluated and for what purpose. However, they do seem useful as general guides to what an ideal set of standards might look like.

#### The Nature of Evaluative Research in Education

In the early chapters of the thesis it was suggested that there has been some confusion over the nature of evaluation and evaluative research in education. Now that a model for evaluation has been proposed, it is possible to attempt to say what the distinctive features of evaluative research might be.

In the proposed model of evaluation, heavy emphasis has been placed on the need to identify or construct explicit standards, particularly since the professional evaluator usually works with

clients, whose standards may not be explicit. Within the literature, and in terms of the activities of many 'evaluators', the need for explicit standards has not always been recognised. As a result, many 'evaluations' have in fact gone no further than description. At the same time, there has been relatively little concern with methods for arriving at statements of standards. Thus Provus writes that "the profession has been so preoccupied with measurement problems that the more basic meaning of criteria in evaluation methodology has been overlooked", and that:

"the explication of a suitable standard at each stage of evaluation constitutes the most difficult part of the evaluation process in a pluralistic society. To date the literature has hardly touched on this critical aspect."<sup>24</sup>

Similarly Stake reports that "most evaluators have chosen not to judge"<sup>25</sup>, and that "most writers do not include judging the worth of alternative objectives and identifying standards as one of the evaluator's jobs."<sup>26</sup> Moreover;

"Manuals and guidelines ..... for project evaluation typically call for gathering statements of objectives without reference to their value loadings. They require no attention to priorities and standards".<sup>27</sup>

One of the results of this de-emphasis on standards has been that 'evaluators' tend to have described rather than evaluated, and to that extent have given some grounds to clients' seemingly widespread discontent. Thus Stake says:



"Michael Scriven has charged evaluators with responsibility for [passing] judgment upon the merit of an educational practice. (Note that he has urged the evaluator to do what the educator has expected the evaluator to be doing.)"<sup>28</sup>

Also that:

"The countenance of evaluation beheld by the educator is not the same one beheld by the specialist in evaluation. The specialist sees himself as a 'describer', one who describes aptitudes, environments and accomplishments. The teacher and school administrator, on the other hand, expect an evaluator to grade something or someone as to merit".<sup>29</sup>

Fortunate indeed, perhaps, that 'evaluators' are not liable under the Trades Descriptions Act!

The importance of standards for evaluation has, of course, been recognised by some practitioners, such as Scriven and Stake. For others, however, the need for standards seems to have been realised 'after the event'. The problem sometimes emerges in the following terms:

"The logic is simple, having adumbrated our aims in behavioural terms, and having constructed measurement instruments to assess whether the aims have been achieved, the test results provide us with the evidence. But difficult problems remain. What is 'success'? Do we expect all pupils to score 100% on our mastery test? Or can we be satisfied with 95 or 90 per cent? .....

The odds are that we find that objective X has an overall 'success' rate of 75 per cent, objective Y one of 60 per cent and objective Z one of 20 per cent."<sup>30</sup>

Weiss writes in similar vein:

"Once the goals are set, the next question is how much progress toward the goal marks success. Suppose a vocational program enrolls 400, graduates 200, places 100 on jobs of whom 50 are still working three months later. Is this success? Would 100 be success? 200? 25?"<sup>31</sup>

The dilemma for the evaluator in standardless situations is well expressed by Stake:

"Many evaluators feel that they are not capable of perceiving, as they think a judge should, the unidimensional value of alternative programs. They anticipate a dilemma such as Curriculum I resulting in three skills and ten understandings and Curriculum II resulting in four skills and eight understandings. They are reluctant to judge that gaining one skill is worth losing two understandings."<sup>32</sup>

Different practitioners have taken different positions in respect of the question of standards. Scriven<sup>33</sup> says that there is no evaluation without judgement, and that the evaluator is in the best position to judge. Stake seems to emphasise the need to collect data relevant to judgement, which he calls 'judgment data'. Others, such as Wiseman and Pidgeon, see the evaluators task as to report data, leaving the determination of standards to the clients or user. Evaluators have tended to recognise that passing judgement involves a certain amount of responsibility, and have naturally been reluctant

to judge where to do so seems to involve applying their own standards. They do not feel easy in their minds that they should be the ones to say what counts as 'success' or 'failure'.

In the writer's view, Scriven is correct to point out that to describe is not to evaluate, and that evaluators might reasonably be expected to produce evaluative statements about the subject they are evaluating. However, it would also seem that in formulating these judgements, the evaluator should not be left in the position of having to apply his own standards. What needs, perhaps, to be done is to identify the standards of clients, however diverse such a group might be. Indeed it is this task which seems to give evaluative research its distinctive character, the identification of the standards to be applied in the evaluation.

Various writers, such as Cherns and Clark,<sup>34</sup> have put forward typologies of research. Constructing such typologies is not an easy matter, and whatever criteria are suggested, it is usually possible to find examples which do not seem to fit.

Cherns lists four types of research in the social sciences, distinguished by their problem orientation and the typical channels of diffusion of their results. Thus he refers to:

Pure Basic Research:- oriented to dealing with theoretical problems within a discipline; results diffused largely through academic publications.

Basic Objective Research:- oriented to studying problems arising in a field of application but not at prescribing solutions; results diffused through academic and professional journals and via specialist teaching.

Operational Research:- oriented to tackling ongoing problems within an operational framework; diffusion of results direct to organisational administrators and decision-makers.

Action Research:- involving introducing and observing planned change; continuous feedback to organisational policy-makers and decision-makers.

It can be seen that Cherns does not refer to evaluative research as a separate type of research. Clark, on the other hand, does. His typology is developed from that given by Cherns. He suggests three dimensions for classifying research; orientation to practical or theoretical problems, dominant channel of diffusion learned journals or reports to sponsors, predominantly single or multiple audiences for results. Five main types of research are identified; Pure Basic, Basic Objective, Evaluation, Applied, and Action. Evaluation research is characterised as oriented to practical problems, diffusing results mainly within the sponsoring enterprise, and having an audience including both members of the sponsoring organisation and other organisational scientists. Typically, says Clark, such research has been concerned with assessing the effectiveness of change programmes within enterprises, and social programmes aimed at tackling problems such as poverty and educational disadvantage.

This characterisation of evaluation research is do doubt adequate for its originator's purposes. It may, however, be possible to provide a formulation which draws out more strongly the particular character of evaluation as the writer sees it. We have already suggested that the evaluative process usually involves two aspects, standard setting or identification and the acquisition of information. Normally, it would seem that in terms of what evaluators have actually done, evaluation has been equated with the acquisition of information. Evaluative research has thus been characterised as the kind of research which aims to describe some state of affairs so that judgements may be made about that state of affairs. This is perfectly reasonable. The difficulty is that since almost any information can be used as a basis for making judgements by someone, it seems that any research might be called evaluative. We can be more specific by saying that evaluative research is oriented to the collection of information which is known to be intended directly to inform judgement. Yet as a research activity, there again seems to be no special feature which marks it as a distinct form of research. It would seem, for instance, that the Civil Servants who collate figures on the number of unemployed, or the trade figures, are engaging in evaluative research, since it is certain that the Minister, and others, will be basing their judgements of the 'unsatisfactoriness' of these phenomena on such information.

In the writer's view, one way of giving a strong characterisation to evaluative research is to say that one of the objects of the research is the identification of standards. That is, that part of the research activity involves obtaining information on the subject that is being evaluated, and part on identifying or creating the standards which are to be applied to that subject. The actual business of formulating evaluative statements is not itself a research activity. It is more like the application of an algebraic formula to a given set of data. Similarly, the methods and techniques used to obtain information about the evaluative subject are similar to those used in any social science research activity. Indeed, according to Suchman, "evaluative research has no special methodology of its own."<sup>35</sup> What does seem to be distinctive is that the evaluative researcher has to investigate the standards which are to be applied to the subject as well as to generate information about the subject itself.

It is very difficult, and perhaps neither very useful or desirable, to make rigid distinctions about what is or is not to count as evaluative research. As has been shown earlier, there are varying degrees of 'sophistication' within standards which have implications for the role of the evaluator. It is nevertheless useful to have a kind of ideal type model of the activities involved in evaluative research and of the role of the evaluator. Thus for us, evaluative research involves an investigation of the state of the evaluative subject and an investigation into the standards to be applied, the relevant population in the latter case

being clients. Evaluation involves the application of standards to the information generated about the evaluative subject.

REFERENCES

1. Urmson (1950), p.146.
2. M.C. Wittrock, 'The Evaluation of Instruction: Cause and Effect Relations in Naturalistic Data', in M.C. Wittrock and D.E. Wiley, eds., The Evaluation of Instruction: Issues and Problems, New York 1970, pp.6-7.
3. Suchman, p.31.
4. The influence of Meehan's work on the views put forward here will be apparent. See E.J. Meehan, Value Judgment and Social Science, Illinois 1969.
5. Ibid., p.99. Meehan refers specifically to human attributes.
6. Urmson (1968), Chapter Five.
7. It is impossible to do full justice to Hartman's complex analysis without a lengthy exposition. It may be that the writer has misinterpreted Hartman on this point. The reader is therefore advised to consult Hartman's text for himself.
8. Thomas, 'Success and Failure'.
9. Meehan, p.115 ff.
10. Emmet, p.126.
11. See, for example, R.L. Ackoff and M.W. Sasieni, Fundamentals of Operations Research, New York 1968.
12. Urmson (1950).



13. A. Kaufman, The Science of Decision-Making, London 1968.  
M.J. Sargeaunt, Operational Research for Management,  
London 1965.
14. E. Herzog, Some Guide Lines for Evaluative Research,  
Washington 1959, p.17. Cited in Suchman, p.54.
15. R. Harré and P.F. Secord, The Explanation of Social Behaviour,  
Oxford 1972, pp. 302-13.
16. G. Vickers, 'Values, Norms and Policies', Policy Sciences,  
4:1(1973).
17. Coombs, p.3; Urmson (1968), Chapter Five.
18. "Sir Thomas Beecham once said of a camel which made a mess  
on the stage at a performance of Aida which dissatisfied him,  
that as an actor it was mediocre, but as a critic superb".  
Quoted in Urmson (1968), p.103.
19. R.E. Stake, 'Objectives, Priorities and Other Judgment Data',  
Review of Educational Research, 40 (1970), p.182.
20. "While it is no doubt true that better information makes for  
better decisions, it is equally true that however good or  
complete the information available, it will be evaluated  
in terms of the policy-maker's theories and values, and  
that mistaken theories and inappropriate values figure  
as largely in bad decisions as does inadequate information."  
A. Cherns, 'Social Sciences and Policy', in A. Cherns,  
R. Sinclair and W.I. Jenkins, Social Science and Government,  
London 1972, p.19.

21. Davis writes, in a discussion of superempirical ends:  
"Since such ends visualize a future state-of-affairs in the superempirical world, it is impossible to prove or disprove in logico-empirical (scientific) terms that any means chosen is adequate ..... with the result that rationality becomes irrelevant and the action is non-rational in character." K. Davis, Human Society, New York 1966, p.128.

Parsons says that "Action is rational insofar as it pursues ends possible within the conditions of the situation, and by means which, among those available to the actor, are intrinsically best adopted to the end for reasons understandable and verifiable by positive empirical science." T. Parsons, The Structure of Social Action, New York 1937, p.58.

Hartman says: "By 'rational' we mean the capacity to combine concepts with objects, which is really the capacity to find one's way in this world by representing it to oneself, that is, by giving names to material objects and interrelating the names". Hartman, pp.116-17.

22. Coombs, pp. 18-19.
23. Sanders and Cunningham, pp. 224-26.
24. Provus, pp. 163, 375.
25. Stake (1967), p. 526.
26. Stake (1970), p. 186.
27. Ibid., p. 187.

28. Stake (1967), p. 526.
29. Ibid., p. 525.
30. Wiseman and Pidgeon, pp. 86-87.
31. C.H. Weiss, Evaluation Research: Methods of Assessing Program Effectiveness, Englewood Cliffs, New Jersey 1972, p. 32.
32. Stake (1967), p. 527.
33. Scriven (1967).
34. Cherns, p. 29; Clark, pp. 6-25.
35. Suchman, p. 81.

## CHAPTER SIX

### EVALUATIVE DISPUTES

In an earlier chapter, it was suggested that one of the important differences between evaluation and description is the relative openness of evaluative terms to the specification of different criteria of employment. It was noted that for this reason, evaluative statements tend to be ambiguous. We have also seen that there are often ample grounds for formulating standards which lead to different evaluations of the same subject. It is therefore likely that evaluative statements are often going to be a focus of dispute, since one of their inherent properties is openness to argument. It seems that there is no guarantee that such disputes can be resolved, but it is worth trying to identify the more fundamental and more superficial levels of disagreement.

Since the model of evaluation proposed here regards evaluation as involving the application of evaluative standards to information, we can see disputes as arising from disagreements within either or both of these areas. But before pursuing this, I should like to quote an amusing, if rather trivial, example of an evaluative dispute.

The Guardian newspaper recently reported as follows:

"The Department of the Environment, responsible for granting contracts to motorway caterers, yesterday published results of a survey indicating that people are satisfied with the amenities offered.

This was less of a surprise to connoisseurs of human nature than it was to motoring and other organisations who still say that facilities are inadequate.

The Department employed a private research organisation to ask 3259 people the degree of their satisfaction or dissatisfaction with the facilities.

Fifty-five per cent of private motorists and sixty-nine per cent of commercial motorists used the catering facilities, and more than seventy per cent of both were either 'very satisfied' or 'quite satisfied' with the quality, flavour and presentation of what they consumed. Twenty-seven per cent of the commercial users were dissatisfied; nineteen per cent of private users.

When it came to hygiene, tidiness and pleasantness of surroundings, even more people were satisfied - over eighty per cent - and only ten per cent were dissatisfied. The opinions of the service were about the same.

The people sampled did not spend much. Sixty-five per cent of private users and seventy-six per cent of commercial parted with less than forty-one pence. Only nineteen per cent of private users and twenty per cent of commercial users said they would have been prepared to pay more for a meal of their choice, cooked individually and of higher quality than the one they actually had.

The results were beaten like a tough steak by people whose business is assessing food. The Egon Ronay Organisation, which last year detected in motorway

area catering 'a slight deterioration on a generally low standard', said that 'we are unfortunately not surprised the public is satisfied'.

The results could have been distorted by the fact that hundreds of commercial drivers deliberately turned off the motorway to eat because they didn't like the facilities provided. 'They use hundreds of places off the motorway as much as they possibly can' the organisation said.

The AA, which two years ago complained of the 'appalling and tasteless monotony of the quick snacks', said that its view had not changed much since.

'The number of complaints we have received from members about the standard of catering has dropped off a little bit. But we still feel the overall standard of motorway catering does leave a great deal to be desired. We told the Department of the Environment that we hoped there would be an announcement before long, and presumably this is it'.

Mr. Christopher Meakin, director of home affairs of the Association of British Chambers of Commerce, said:

'This survey is not at all convincing .... Having been connected with the survey business, I know you can get any result in advance by the way you wrap your questions up'. Simply asking people whether they were satisfied or dissatisfied was not the best way of evaluating amenities.

It was time that big contractors ceased to have the monopoly in motorway catering. 'There is nothing better than Momma and Poppa service, and the further you get from that, the worse it gets.

I would like to see the sort of people you see in Charlotte Street and Beauchamp Place given a

chance to get in with some lasagne and spaghetti'.

The Department did admit yesterday that it was studying certain 'minority complaints' thrown up (if the phrase may be excused in this context) by the survey. These were mostly about prices, or coffee being too strong or too weak, too hot or too cold."<sup>1</sup>

The specific research which is referred to in this article is not of great importance here, but the report does illustrate some interesting features which are common to evaluative disputes in general. The dispute in this case is over the evaluative status of motorway catering amenities. The Department of the Environment claims that they are 'adequate', whereas other interested parties are claiming that they are not. Disagreement over both standards and information appear to be present.

Firstly, there is a dispute over standards. The Department presumably defines the services as 'adequate' in terms of a majority of respondents indicating 'satisfaction' with various aspects of the service. The AA, however, refers to its feeling that "the overall standard of motorway catering does leave a great deal to be desired", and Mr. Meakin says that "there is nothing better than Momma and Poppa service, and the further you get from that, the worse it gets". Although the standards of the critics are not clearly stated here, it is apparent that each party is applying a different set of standards.

Secondly, the status of the evidence collected by the Department is questioned. "The results could have been distorted

by the fact that hundreds of commercial drivers deliberately turned off the motorway to eat because they didn't like the facilities provided"; "Having been connected with the survey business, I know you can get any result in advance by the way you wrap your questions up. Simply asking people whether they were satisfied or dissatisfied was not the best way of evaluating amenities". Some of the critics seem to be saying that even if the Department's standards were sound, the amenities do not meet the criteria of adequacy, or that the available information makes it impossible to say whether the amenities are adequate or not.

The importance of evaluative disputes is most clearly seen where the disputants need to agree before actions can be carried out. Clearly, if there is a dispute over whether some state of affairs is 'satisfactory' or not, there is also likely to be a dispute over whether anything should be done about it, and if so what should be done about it. It may well be that no resolution is possible, but it also seems that often the grounds for disputes are insufficiently explored.

Looking firstly at disputes on the information side of evaluation, it seems that these are likely to consist of opposed factual claims about the subject being evaluated. In such a situation, the disputants might ask each other about the sources of their information and the methods they have used to arrive at their factual claims. Supposing that these turn out to be the same for both parties, then one might claim that the other had somehow misapplied the method, for example 'miscounting' a set of



items. If this is not accepted, then the parties could jointly reapply the methods they have used and hopefully arrive at an identical result.

A second possibility is that the parties have used different methods to arrive at a result, and that each method has been correctly applied. In this situation, one party may attempt to resolve the dispute by trying to convince the other that his method of arriving at a result has the greater claim to acceptability. In other words, the argument becomes one concerned with the validity of the methods used to arrive at factual claims.

As the methods of the sciences are generally regarded as the ones which produce the highest quality empirical data, everyday methods of acquiring information are particularly vulnerable to attacks on grounds of invalidity when research data are on hand. It might be thought, however, that disputes which consist of opposed factual claims based on everyday methods will be automatically resolved by instituting a research investigation to establish the facts. Although it is true that the fact that findings have been generated by research work tends to be used as a basis for claiming their overriding authority, it is also the case that the social science methods typical of formal evaluation are themselves open to questioning on validity grounds. As the example above shows, even when everyday methods are replaced by more sophisticated techniques of data-collection, the findings are still open to argument.

The example of the motorway amenities reveals a dispute about what is to count as evidence of customers' satisfaction. Reference

is made to being able to get any result from a survey, and although one might want to argue with this assertion, it certainly seems to be the case that in the social science field there is often dispute about the validity of 'evidence'.

An example from Open University experience illustrates the point. Students taking Open University course have to complete a number of written assignments which are sent to tutors for marking through the post. Each assignment is accompanied by a form, on which the student is asked to record, among other things, the 'time taken' to complete the assignment. The idea was that this information would enable assignments that were taking 'too long' to be identified, so that modifications could be made for future years. Yet even if there had been agreement on what the number of hours which would count as 'too long', the validity of the information given by students could easily be questioned. Students were not given a definition of what was to count as 'time taken'. Was it the time to write the assignment? Perhaps it was the time to write and prepare it. Or perhaps it was the time taken just to prepare it. And how valid would students estimates of time be anyway? Perhaps they thought they would be penalised if they indicated 'too many' hours. Perhaps they thought they would get more credit from their tutor if they showed they had spent a 'long time' on the assignment. Possibly they put an arbitrary figure in out of habit. In this way, seemingly straightforward data can be very easily undermined. We might also suggest that an attack on the validity of information is quite likely to be the strategy of those who are committed to particular standards, but who do not like the look of

the evaluative statement which emerges from the standards to which they have assented.

Shipman<sup>2</sup> has written at some length about these kinds of limitations on the capacity of social science research to produce unambiguous evidence, and it is worth quoting one or two extracts.

For instance:

"But in all social science there is a pressure to produce results and produce them in unambiguous form. Foundations, Departments, and businesses giving money for research expect results as a sign that their money has been well spent. The continuation of research in a department and the employment of those involved depend on producing the goods. Yet certainty and clarity are often impossible in the messy arena of everyday life. The ambiguous results and imperfect methods are cleaned up for public consumption. In the market-place, as elsewhere, it is all things bright and beautiful that sell. The honest remain not only poor but unpublished."<sup>3</sup>

"The conclusion to be drawn from the fate of social science predictions is that there is little chance that complex topical problems are open to solution through available methods of research."<sup>4</sup>

"All the controversies of this book confirm this inability of social science to provide decisive answers to crucial questions."<sup>5</sup>

The problem is particularly acute when the subjects of evaluation are alternative sources of action. Since these will often be evaluated in respect of their ability to produce a preferred outcome, the kind of indisputable information that is

required for the execution of the evaluation concerns the relations between the variables which will be manipulated as a course of action is ultimately intended to influence. The limited capacity of the social sciences to produce any information about many relations, and only limited information about particular relations in particular circumstances, means that there is ample scope for information to be a source of dispute in such evaluative situations.

The recent case of Britain's decision on whether to remain a member of the Common Market provides a good illustration of the problem. An examination of the pamphlets put out by the 'Pro' and 'Anti' organisations reveals that there is disagreement on nearly every factual point.<sup>6</sup> Presumably, the social sciences are simply not in a position to predict what the consequences of staying in or coming out would be, although some possibilities can be ruled out. As a result, there is ample scope for disagreement about the state of affairs that will ensue from either course of action.

The magnitude of the informational problem will obviously vary from case to case, but it seems that where major changes are considered in large and complex organisations, it is at least likely to be difficult to establish what the effects of those changes will be. As Meehan says:

"If no more can be done than to drive the discussion of values to a point where the empirical dimensions of the argument have been agreed and the sole remaining questions have to do with the order of priorities to be assigned to different variables in in different situations, that would be a cause for



Order of set of  
significant descriptions:

+ order identical

- no elements share the same place in  
the respective orders

Evaluative term set:

+ participants' evaluative term sets  
are identical

- participants' evaluative term sets  
share no members

Evaluative function:

+ participants' evaluative functions  
identical

- participants' evaluative functions  
not identical

In each case it is assumed that there is prior agreement on the definition of the evaluative subject, so that a minimum condition for being able to agree or disagree about instances of the subject is that there is agreement on what subject is being talked about.

Situation 1 is the case of complete agreement on all the components of the standards, which necessitates agreement on the evaluative status of any instance. It can be noted, however, that even agreement on the whole evaluative scheme may not entail agreement on its justification. We could agree that a six-room house is satisfactory and houses with all other numbers of rooms unsatisfactory, and hence produce identical evaluative statements for any instance of a house. Yet my interest in rooms may stem from my family circumstances, whereas yours stems from a simple preference for that sort of house against all others.

Figure 10: Configurations of Agreement and Dispute  
within Standards

	1	2	3	4	5	6	7
Criterion variables	+	+	+	+	+	-	-
Order of set of significant descriptions	+	+	+	-	-	-	-
Evaluative term set	+	+	-	+	-	+	-
Evaluative function	+	-	-	-	-	-	-

In situation 2, everything is agreed bar the evaluative function. Such disagreement can only occur when the number of evaluative terms is less than the number of elements in the set of significant descriptions. If they are the same in number, then, as has been shown earlier, there is only one way of defining the function. Where this situation does arise, there will still be agreement on the evaluative statuses of the elements falling first and last in the order of significant descriptions, but disagreement over some of the other elements.

In situation 3, criterion variables and the order of significant descriptions are agreed, but as the evaluative term sets are non-overlapping, the evaluative functions cannot be identical. In this case, there is no agreement on the evaluative status of any element. It may be that in such circumstances, the only real point of disagreement is a verbal quibble over the evaluative terms, the function being basically the same. Agreement can then be easily secured by coming to some compromise set of terms. On the other hand, the different evaluative term sets may indicate that the participants are evaluating from different points of view.

In situation 4, the different orders of the set of significant descriptions prevents identity of the evaluative functions. There may be agreement on the evaluative statuses of some of the elements but not on all of them.

Situation 5 is similar to 4, but now there is disagreement on the evaluative term set. This eliminates the possibility of



agreeing on the evaluative status of any item.

Situation 6 constitutes disagreement on every component except the evaluative term set. Since the participants hold no criterion variables in common, the orders of the significant descriptions must be non-identical since the set contain different elements. Similarly the evaluative functions must be different since the elements in the domain are different. Yet were there an agreed set of evaluative terms, it would still be possible to agree the evaluative status of every instance evaluated. Suppose, for instance, that the set of items to be evaluated is partitioned into two equivalence classes, each participant using a different two-valued criterion variable. They each define a function between their criterion variable and a shared evaluative term set. If cases of the evaluative subject then always share the two values of the criterion variables which correspond to one evaluative term, and also the other, each instance will be attributed the same evaluative status by the participants.

Of course the agreement here is a very tenuous one in that the only component of the evaluative scheme which is shared is the evaluative term set.

Situation 7 represents disagreement on every component of the evaluative scheme, and hence of disagreement about the evaluative status of every instance of the subject.

Looked at from another point of view, we can see that agreeing on the evaluative status of an item might lead us to

imply at best that the whole evaluative scheme is agreed, or at worst that the only component actually agreed is the evaluative term set. If there is disagreement over the evaluation of an item, this might imply at best disagreement over the evaluative term set, or at worst disagreement over the whole evaluative scheme.

In any event, it would seem that we always need to be wary of assuming that we share criteria even when we agree on the evaluation of an item, and similarly wary in assuming fundamental disagreement when we disagree over the evaluation of an item.

Given that to obtain a full understanding of what a person means by an evaluative statement, we need to know what his criteria are, it seems that we are likely to be uncertain about this until they have been made explicit. Unfortunately, we cannot assume that because we agree on an item's evaluative status that we share the same criteria; nor because we disagree that we do not.

The process of externalising evaluative schemes is thus the first step to resolving disagreement, indeed to the identification of disagreement. Doing this may not lead to any resolution of whatever conflict is present, but in the process of carrying it out the understanding of participants of both their own and others evaluations is likely to increase.

In some situations the externalisation of evaluative schemes may present no problem. Urmson's apple grader can probably say what his standards are, if he has not become so adept at applying

them that he has lost the ability to express them. Even if he has, he can always refer to the officially printed specification.

From the point of view of professional evaluation, it is necessary to secure explicit evaluative schemes for two reasons. Firstly, it seems desirable that those who are collectively responsible for maintaining some state of affairs at a level which conforms to certain standards should be able to agree the content of the standards. There does not seem to be much point in agreeing to ensure a satisfactory state of affairs without also agreeing what constitutes a satisfactory state of affairs. This may never be an easy thing to do, but it might be systematically approached via attempting to unpack implicit evaluative schemes, or construct new ones via the approach already suggested.

Explicit evaluative schemes are also necessary to the professional evaluator because he is usually asked to evaluate on the client's behalf, that is by proxy. He can clearly only do this if the client is able to make his evaluative scheme explicit, although as has been shown earlier, how much of the evaluative scheme it is necessary, or possible, to define will depend on the circumstances. It is no wonder that professional evaluators are sometimes in a quandry about what judgement to pass on the subject of the evaluation. If they have not secured a statement of the client's standards, their only recourse is either to simply present the data or to try to generate standards out of their own heads. Since they often do not feel qualified to do this, they face a dilemma if they really are expected to

pronounce on 'satisfactoriness', 'success' or whatever. It seems to the writer that it is the client's job to specify standards, albeit with the help of the evaluator.

Within the literature on educational evaluation there is some disagreement on the question of whether the evaluator should be concerned with making judgements, that is generating and applying an evaluative function in respect of the data collected. Stake, for instance, writes:

"Description is one thing, judgment is another. Most evaluation specialists have chosen not to judge. But in his recent Methodology of Evaluation, Michael Scriven has charged evaluators with the responsibility for passing judgment upon the merit of an educational practice .... Scriven's position is that there is no evaluation until judgment has been passed, and by his reckoning the evaluator is best qualified to judge."<sup>8</sup>

Wiseman and Pidgeon, however, take the reverse view in their book on curriculum evaluation:

"If curriculum A achieves most success with objectives 1, 3, 5 and 7; curriculum B with objectives 2, 4, 6 and 8, then what? The potential user must judge for himself his aim priorities, choosing in the light of his own philosophy and his own imperatives. He is faced with a professional value-judgement, the responsibility for which cannot be shrugged off of on to the shoulders of an unknown evaluator".<sup>9</sup>

"Only the teacher can decide whether a particular average score on a test measuring a particular objective is high enough to give cause for satisfaction or inspires doubts on the effectiveness of the instruction given or the methods used."<sup>10</sup>

In situations where the evaluator and the decision-maker are the same person, the problem of who should set standards does not arise. Normally, however, the evaluator and the client/decision-maker are separate individuals, so that there is a question as to who is going to set standards.

One solution is for the evaluator to take on the role simply of researcher, which according to Stake (see above) is what most evaluators have done. In this case, the evaluator describes and the actual business of judgement is left to the client. Yet even here, the client must indicate which variables he is interested in, his criterion variables, if the evaluator is to be able to provide relevant information. In this event, however, the 'evaluator' is not evaluating; he is providing information which others will use in the process of evaluation. He is thus doing very little more than any applied researcher, and one might doubt the necessity to refer to the task as 'evaluation'.

One of the difficulties with this approach is that the client is quite likely to be in some difficulty in formulating explicit standards. If the aim of professional evaluation is to improve the quality of value judgements, then it would seem that part of the evaluation specialist's task will involve helping the client to arrive at an adequate specification of standards. This

seems particularly pertinent if, as Coombs says, "a great many people, perhaps most, do not always think rigorously in making value decisions."<sup>11</sup>

The distinctive aspect of the evaluator's role therefore seems to be that an important part of his task is the identification of standards. It is this which seems to the writer to give evaluation research its claim to be a distinct form of activity.

There will be situations, however, where it is by no means clear who the client is. There will then be a problem about whose standards should be used, or who should be consulted about standards. This is, of course, not unrelated to the problem of deciding on a point of view to inform the evaluation.

Some consideration has been given to this topic by evaluation practitioners. Provus, for instance, says that under his model "only the staff (which includes administrators and practitioners) has the power to define program standards", although he believes that most decision-makers "include the values of public school constituents in their criteria for selecting and modeling [sic] program standards."<sup>12</sup>

Weiss approaches the problem in terms of identifying users. Possible users of the evaluation include funding organisations, national agencies, local agencies, project directors, direct service staff, the group at which the activity being evaluated is aimed, and scholars in the disciplines and the professions. On the one hand

the evaluator may be "on the staff of some organisation ... and he does the job assigned to him"; or he may be outside the organisation and able to "negotiate the purpose and focus of the study." The evaluator is always free to utilise his own values, for "it is important that the evaluator be able to live with the study, its uses, and his conscience at the same time."<sup>13</sup> Beyond this, one must consider the decisions to be made and who is interested.

Obviously no hard and fast rules can be given as to whose standards should be used in an evaluation, since so much depends on the context of the evaluative problem. We simply note that if it is accepted that an important part of the evaluator's role is the identification of standards, then evaluator's will often be faced with the problem of whose standards, or who to consult about standards.

#### Resolving Evaluative Disputes

According to Mieux, although much has been written on the topic of conflict and conflict resolution, "many of the writings are of questionable relevance to the resolution of value conflicts."<sup>14</sup> Similarly, there is a lack of procedures for tackling the problem of the resolution of such disputes. We shall therefore suggest some ways in which the problem might be approached, in relation to the analysis of the evaluative process already given.

One of the difficulties is that there is no way of determining simply from the evaluative statements whether there is

a conflict or not. The presentation of two statements such as 'x is satisfactory' and 'x is not satisfactory' immediately suggests dispute, but as has been shown above, even when the evaluative statements are identical, the evaluative schemes and justifications may be at variance with each other.

It therefore seems that whatever the match of the evaluative statements, it is often likely to be worth exploring the evaluative schemes which underlie them, which largely involves making these schemes explicit. We have already seen the need for this when the evaluator is evaluating on behalf of a client. In that situation, however, the purpose of the exercise is to enable the evaluator to comprehend his client's standards. At its most extreme it is simply a one-way communication from client to evaluator. A conflict situation is envisaged as arising, however, when there are multiple clients who are collectively responsible for defining standards. Here, the externalisation of each parties standards is a means to the end of identifying whether there is a dispute between them, and if so, the nature of the conflict.

The first step in the process of attempting to identify and resolve evaluative disputes is thus the externalisation of the evaluative schemes of the participants, together with the information about the subject of the evaluation which the participants have used to locate it within the set of significant descriptions. This involves asking each participant to 'fill in' each component of the evaluative scheme. When this is done, the schemes can be compared to identify mismatches.



In general, the sorts of questions that might be asked about the evaluative schemes include:

1. Are the participants evaluating the same subject? Do they share a common definition of the subject?

If evaluative statements are rendered about subjects which are different, there is really no dispute over the evaluation, since the participants are not really playing the same game. They are not evaluating the same thing differently, but different things. Such a situation can easily occur when the terms used to refer to the subject are open to several interpretations. Consider the case of evaluating the alternatives of withdrawal or non-withdrawal of America from the Vietnam war. Assuming that the evaluative subject set, withdrawal, non-withdrawal, is to be described in terms of consequences, it clearly makes a great deal of difference whether withdrawal means withdrawal of troops, civilians, financial aid, equipment, moral support, or any combination of these.

2. Are the participants using the same point of view?

As noted earlier, evaluative subjects can often be evaluated from many different points of view. The use of different points of view may not be apparent from the schemes themselves, but will presumably become apparent as the process of externalisation proceeds.

3. Are the participants using different criterion variables?

4. Do the participants share the order of the set of significant descriptions.
5. Are the participants using the same evaluative terms?
6. Do participants share the same evaluative function?
7. Do participants share common information about the state of the evaluative subject?
8. Do participants share similar standards of validity in respect of the information?

If these questions can be answered, a great deal has already been learned. However, this process simply reveals whether there is a dispute and what kind of dispute it is. It does not of itself suggest how such disputes might be resolved. Some suggestions can be made, such as:

1. Encourage the exploration of possible definitions of the evaluative subject, in the light of the purpose and context of the evaluation.
2. Seek acceptable grounds for the inclusion and exclusion of competing points of view. For instance, doctors evaluating contraceptive techniques might try to justify the exclusion of their religious views on the grounds that as doctors their sole concern is with the medical point of view.
3. Identify reasons for using different criterion variables. Try to identify the most important and see if agreement can be reached on at least the relevance of these.

4. Identify the magnitude of the differences in the order of the set of significant descriptions. Minimal adjustments may be all that is necessary to reach agreement.
5. Explore the significance of having different sets of evaluative terms. If the differences represent no more than a terminological quibble, select one set or produce a compromise set.
6. Identify the magnitude of the differences over the evaluative function. Minor adjustments may produce agreement.
7. Make information used by participants to make evaluations available to other participants. Encourage participants to consider the relevance of all available information.
8. Clarify bases for claiming validity for information. Refer where possible to 'authoritative' standards of validity. Seek additional information to support or refute information claims.

It should be clear, however, that no procedure can guarantee agreement over evaluation. Perhaps the easiest situation to handle is that where standards are agreed, but the evaluative statement has been formulated in respect of beliefs about the evaluative subject which happen to be easily confirmed or refuted by reference to some established body of knowledge or by the application of some routine data-gathering procedure. Much more difficult is where the dispute arises because of a conflict at the level of Values, preferences and purposes. Such disputes may be irresolvable, other than by arbitrary means. If, for example, A prefers x to y, and B prefers y to x, it seems that there is very little to be done

about it. It may thus be that we shall often fail to resolve evaluative disputes, in the sense of getting complete agreement on every aspect of the evaluative scheme. We may always have a plurality of value judgements. Yet the exploration of evaluative disputes and differences can nevertheless be regarded as valuable, for at least two reasons.

Firstly, the exploration of evaluative schemes is likely to at least increase the participants understanding and appreciation of each others views. It might be possible to appreciate why there is disagreement, even if it is not possible to resolve it.<sup>15</sup>

Secondly, a systematic uncovering of evaluative schemes exposes the implicit reasoning of the judge to the scrutiny of others, and helps to ensure that our often poorly formed evaluations are put to some sort of test. At the same time, the sharing of different points of view and of information may help to improve the quality and quantity of the reflections which lead to them, and help to broaden the basis upon which an individual's evaluations are made.

On the debit side, if it can be called a debit, is the possibility that in systematically exploring our evaluative schemes, we may discover that we are much more confused, more uncertain and less 'rational' than we thought. Perhaps this is the price of enlightenment. As Wright-Mills has said:

"But when there are values so firmly and consistently held by genuinely conflicting interests that the

conflict cannot be resolved by logical analysis and factual investigation, then the role of reason in human affairs seems at an end. We can clarify the meaning and the consequences of values, we can make them consistent with one another and ascertain their actual priorities, we can surround them with fact - but in the end we may be reduced to mere assertion and counter assertion; then we can only plead or persuade. And at the very end, if the end is reached, moral problems become problems of power, and in the last resort, if the last resort is reached, the final form of power is coercion.

We cannot deduce - Hume's celebrated dictum runs - how we ought to act from what we believe is. Neither can we deduce how anyone else ought to act from how we believe we ought to act. In the end, if the end comes, we just have to beat those who disagree with us over the head; let us hope the end comes seldom. In the meantime, being as reasonable as we are able to be, we ought all to argue."<sup>16</sup>

### The Contentious Context of Evaluative Research

We note in an earlier chapter that evaluative research tends to be undertaken under the auspices of large formal organisations. The case of Open University counselling is clearly set in this kind of context. As we will see, the call for evaluative research arose from a dispute over the desirability of devoting resources to some ends rather than others, and we now intend to argue that evaluative research is typically called for under these conditions, that is conditions of dispute.

In large organisations major decisions tend to be delegated to specialised decision-making bodies, such as committees, which typically consist of members who have particular responsibilities in the organisation. At the Open University, for example, the body which devised the structure of the tuition and counselling services included representatives from the Faculties, the Institute of Educational Technology, the regional organisations, and so on. At the same time, the specialist nature of the research process tends to mean that there is a distinct department (evaluation unit, research department, etc.) which serves the organisation as a whole. Typically we find a situation where the decision-making body requests or calls in the research department to tackle a particular problem. That is to say that the decision-making body requests or calls in the research department to tackle a particular problem. That is to say that the decision-making body has to decide, or choose, to initiate or authorise research work.

Most decisions involve an element of evaluation. It would, of course, be possible to make a decision on whether or not to activate research work by random methods, such as the coin toss, but the decision-making milieu of the formal organisation tends to require reasoned decisions. A decision to initiate evaluative research thus presupposes an evaluation of the alternatives of initiating or not initiating it. Earlier we examined the distinction between informal and formal evaluation, the former being the kind which we typically use in everyday decision situations. It seems, therefore, that any decision to initiate evaluative research must involve an informal evaluation of a subject state of affairs and

the alternative courses of action of initiating or not initiating research.

Let us introduce some simplifying assumptions. We will assume that the decision-making body meets to decide whether or not some state of affairs for which it is responsible is or is not satisfactory. Let us also assume that the members of the body have real agreement on the standards to be applied. Further, let us assume that it sees evaluative research as purely an information providing activity, and that it knows that it has the option of putting researchers to work to increase the information it has on the state of affairs. The question then is; under what conditions is the body likely to actually call for research?

The decision-making body as we have described it, is seen as a collection of individuals, but we can also apply the argument to a body which consists of just one person. In either case, we assume that the body intends to resolve the question of whether some state of affairs is satisfactory. We also assume that the body has some procedure for arriving at an answer which is to be taken as the answer of the body, such as the majority vote, the chairman's decision, etc.

As the body begins to tackle the question we suggest that its members will apply the agreed standards to whatever information is at hand about the state of affairs. In a vague and intuitive way it will offer preliminary answers, preliminary evaluative judgements. It will in fact indulge in informal evaluation, in so far as it uses data derived from everyday observation, hearsay, and so on. The

outcome of this informal evaluation will take three main forms:

1. The body agrees that the state of affairs is satisfactory
2. The body agrees that the state of affairs is unsatisfactory
3. The body is uncertain about the satisfactoriness of the state of affairs.

In the case of a multi-member body, uncertainty arises, given agreement on standards, when different members make opposed factual claims. In the case of the individual decision-maker, it is as if he has two people in his own head making opposed factual claims. In each case, the body is in dispute with itself. Note that even if all the members of a body were to agree that they were uncertain about the evaluative status of a state of affairs, the body is still in evaluative dispute.

Returning to the three outcomes listed above we suggest the following typical strategies:

1. When a body agrees that some state of affairs is satisfactory it will not call for research (i.e. further descriptions of that state of affairs).

There are a number of reasons for this. One is that in practice there is always an infinity of claims on resources for the purpose of creating states of affairs. Whenever resources are devoted to generating descriptions of states of affairs, they are clearly not being devoted to actually creating the states of affairs themselves. In other words, the more money we devote to determining, say, the



types of problems which Open University students raise with their counsellors, the less money we have to employ counsellors, print books, provide counsellors with training and so on. Since organisations are primarily oriented to doing things, rather than describing what they do or have done, if a state of affairs is agreed to be satisfactory resources are much more likely to be claimed for the purpose of achieving the organisation's goals than for describing more 'thoroughly' that state of affairs.

A second reason is that whenever a body asks for a research generated description, it always opens itself to the possibility that its informal impression is inaccurate. I may feel confident that nine out of ten of my friends admire me and regard this as a satisfactory state of affairs. If I then decide to administer an 'admiration' questionnaire, I at least run the risk of finding that nine out of ten of my friends do not admire me. We can put forward the hypothesis that most people are reluctant to initiate further inquiry into a state of affairs if they believe it has a positive evaluative status. We do not tend to find a doctor who looks at a person and pronounces him healthy then administering a battery of physiological tests. Nor do we find a teacher who is convinced he is teaching well, begging for someone to provide him with objective test data on his students' learning. Nor, we can suggest, would we find the Open University calling for research into counselling if everyone agreed that the counselling service was satisfactory.

An interesting point to notice is that if what we have said is true, it says nothing about whether this tendency is of itself desirable. If we believe that an evaluation should be based on the best quality information, and that this really means data obtained by 'scientific' means, then in principle any and every evaluation should be based on research evidence. It may be that the tendency to accept informal evaluations when they are positive simply means that we are complacent when we should not be. We simply suggest a general tendency for decision-making bodies to behave in this way.

2. When a body agrees that some state of affairs is unsatisfactory, it will not call for research (i.e. further descriptions of that state of affairs).

The reasons for this are related to those mentioned above concerning scarcity of resources. If a body agrees that some state of affairs is unsatisfactory, it is much more likely to devote resources to rectifying that state of affairs rather than to obtaining additional information. If we believe our house is on fire, we are more likely to call the fire brigade than run round the rooms to see which ones are burning. Where things are unsatisfactory there is a pressure to do something about it which does not exist when they are satisfactory.

There is the same possibility as was mentioned above that if resources were devoted to generating descriptions, the informal evaluation that things are unsatisfactory might turn out to be

mistaken. So it appears on the face of it to be worth using resources in that way. The problem is that if the research just confirms that things are as bad as was suspected, money has been spent, time has been lost, and nothing has been done to change the state of affairs. The tendency, we suggest, is for the body to move immediately to a search for solutions. Now at this point we can see another possible occasion for calls for research. However, this would be research into means, and the arguments already given apply. That is, if the body agrees that x is the best thing to do, it will go ahead and do it without asking for a formal evaluation of means; if it agrees that x is not the best thing to do, it will move to the consideration of other alternatives; and if it is uncertain, then see below.

3. If a body is in dispute over whether a state of affairs is satisfactory, it may call for research (i.e. further descriptions of that state of affairs).

In an exemplary case, for the body to be in dispute we find two main possibilities. One is that one member is certain that the state of affairs is satisfactory whilst another is certain it is not. The other is that both members (be they individuals, sub-groups, or psychological individuals) are uncertain. In both cases the body as a whole is uncertain. That is to say that after all decision procedures have been applied (e.g. majority vote, chairman's verdict, etc.) the body cannot say whether the state of affairs is satisfactory or not. For this reason it cannot act to change the state of affairs. There is an impasse.

It is this situation, we suggest, which provides the context for most calls for evaluative research. Evaluative research is called for under conditions of dispute.

It is important to remember the assumptions which we have used as a basis for this argument. In particular we assumed that the body always agrees the evaluative standards to be applied to the situation, so that the only source of dispute was information. In practice dispute may arise because of implicit or explicit disagreements over standards, with or without disagreements over information. In practice, therefore, calls for research may occur when what is required is not better information but also clearer communication of standards. As we have already indicated, disputes in this realm can be handled by resolving procedures rather than research investigations. This leaves us with the business of conducting evaluative research, and this seems most likely to be called for under conditions of evaluative dispute of the kind we have outlined. More precisely, it is most likely to actually be conducted under these conditions, since if the exploration of standards resolves the dispute over standards we are in the position proposed by our assumptions. If it does not resolve this dispute, then the conflict is irresolvable (i.e. no additional information which can be generated by research can lead to resolution). Of course different members within the body may propose different standards, but provided that the body as a whole agrees on which of these to apply, perhaps deciding to apply all of them, then it agrees on standards.

The implications for the conduct of evaluative research will be better appreciated when we have pursued this line of argument further, but we can see immediately that evaluative research becomes cast in the role of conflict resolution. Typically, evaluative research is not undertaken simply to find out whether something meets certain standards. It is carried out to settle arguments.

We have tried to show here that evaluative research arises from dispute. Now we will attempt to make more restricted assumptions about the conditions surrounding calls for formal evaluation. We start from the point we have just established, namely that calls for evaluative research tend to arise when a decision-making body is in dispute over the evaluative status of a state of affairs. In the situation of Open University counselling we find that this was indeed so. We now wish to argue that this uncertainty will rarely be completely balanced, but will tend to one evaluative pole or another. For ease of exposition we will assume that the only evaluative possibilities are 'good' and 'bad'; terms like 'satisfactory' would do equally as well, but 'good' and 'bad' and conveniently short words!

By saying that the uncertainty will be unbalanced we mean that a body will rarely be completely neutral about a state of affairs. Its informal evaluations in the uncertainty situation will tend to be couched in terms such as 'We suspect that this is good, but we're not sure' or 'We think this is probably bad, but we could be wrong'. Since a body usually has responsibility for more than one state of affairs, we tend to find an overall mix of

agreement and dispute over the range of affairs. To make the point clearer, the Open University's Tutorial Board was a large committee which was responsible for many areas such as the counselling service, the face-to-face tuition, the summer schools, and so on.

These constituted different areas for which the Board had a concern.

We can envisage each one of these areas as being informally evaluated, so that if we looked at them all at once, we would find some agreed as good, some agreed as bad, and some in dispute.

The ones that are agreed do not enter into the decision on whether to call for evaluative research. This leaves us with the ones that are disputed. Among these we find some which are suspected bad.

Let us suppose that sufficient resources are available to enable all of the disputed areas to be subjects of research. Even in this situation, we expect that it is the areas which are suspected as bad which are most likely to actually become subjects of research. This is for the same reasons we gave earlier for the non-pursuit of research in certainty situations. In other words:

1. If a body agrees a state of affairs is good -  
no research
2. If a body agrees a state of affairs is bad -  
no research
3. If a body is uncertain, its formal evaluation  
is more likely to be biased positively or  
negatively rather than be neutral
4. Research is more likely to be called for when  
the informal evaluation is biased negatively  
than when it is biased positively.

We are trying to show here that not only will research tend to be called for under conditions of dispute, but also that it will tend to be applied in circumstances when an undesirable state of affairs is suspected. For this reason the body is likely to be anxious about the results of the research, because it expects, on the basis of its informal evaluation, that things are bad.

We can illustrate this further by considering a person who is responsible for two discrete states of affairs or areas (A and B). His overall aim is to have an optimal evaluative configuration, where this consists of evaluative statements derived from informal or formal evaluations or both. If he is certain that both states of affairs are either good or bad he does not call for evaluative research. Let us assume, however, that he is uncertain of one with a positive bias and uncertain of the other with a negative bias. If he only has the resources to apply research to one of these, then the possible outcomes are:

1. the good is confirmed as good; the bad is unchecked - A is known to be good; B is suspected to be bad
2. the good is in fact bad; the bad is unchecked - A is known to be bad; B is suspected to be bad
3. the good is unchecked; the bad is confirmed as bad - A is suspected to be good; B is known to be bad
4. the good is unchecked; the bad is in fact good - A is suspected to be good; B is known to be good.

If the person applies his resources to the state of affairs which is suspected to be good (A), he opens up the possibility that his positive expectations will be overturned, leaving him with an overall negative situation. If, however, he applies his resources to the state of affairs which he suspects is bad (B) the worst that can happen is that his expectation that one of the states of affairs is bad will be confirmed. It is possible, though, that it will turn out to be good, thus leaving him in an overall positive situation.

Although we might agree that dispute within a decision-making body is a frequent concomitant of calls for evaluative research, it is also apparent that the impact of this upon the conduct of evaluative research inquiries will vary according to more specific features of the context. It is one thing to call for the evaluation of the things which are ones own responsibility, and another is to call for the evaluation of someone else's. We often find that organisational decision-making bodies are composed of members who are themselves responsible for the matters with which the body is concerned. In the process of informal evaluation, which must always precede a call for formal evaluation, we find members putting forward their cases for their own areas. Now it would seem that when doing this, members are unlikely to say that they are uncertain as to whether the things they are responsible for are satisfactory or not. It is much more likely that they will have a strong view, which is not only what the organisational culture expects of them, but is also a consequence of the fact that the member is spending his working life in close involvement with the area. He will



probably have conducted countless informal evaluations of how 'things are going', and will have come to the meeting with either the intention of defending his patch against allcomers (i.e. establishing that his area is satisfactory), or of convincing everybody that things are so bad that he must have immediate action.

If a member claims that his own area is satisfactory, the other members may either agree (in which case there is no call for research), or they may claim that it is unsatisfactory. Alternatively they may be uncertain. In either case the body may eventually end in dispute. Since the member is himself certain that his area is satisfactory (that is, as a body he is not in dispute with himself) it follows that he will not call for research. The body, however, is in dispute and it may well call for research. As the focus of dispute is the member's area, and as he is not calling for research, then it must be other members who are calling for it. In other words, when a body calls for research, it tends to be calling for research into an area which the member responsible for that area does not himself see as necessary. If the member claims that his own area is unsatisfactory, the same argument applies.

It would be dangerous to push reasoning of this kind too far. It is no doubt true that calls for evaluation arise under conditions other than these. What we hope to have identified is the modal case in what may be a fairly flat distribution.

In the sort of circumstances we have described, the body ends by calling in the researchers to carry out an investigation into an area of responsibility which the person responsible does not in fact want. He may tolerate the work, even help it along, but he is unlikely to give it his full-hearted consent. Quite possibly he will have put forward quite substantial claims about what is being achieved in an attempt to convince the other members of the body to accept his evaluation. The prospects of research now opens up the possibility that he may have been wrong, and this prospect is hardly likely to be welcomed.

Our conclusions on the contentious context of evaluative research can be summarised as follows:

1. Evaluative research tends to be called for when a decision-making body cannot agree on the evaluative status of some state of affairs. It does not tend to be requested when there is agreement that the state of affairs is either 'satisfactory' or 'unsatisfactory'.
2. When a decision-making body is in dispute over the evaluation of some state of affairs, it is more likely to request research into those which it suspects are 'unsatisfactory' than those which it suspects are 'satisfactory'.
3. Where a decision-making body is concerned with states of affairs which are the responsibility of its own members, the states of affairs which are to be subjected to research tend to be those which are not regarded as in evaluative doubt by the member responsible for them.

This analysis, then, would help to explain why evaluation research is often beset by researcher-client relationship problems. The lukewarm attitude of reluctant client's is understandable. Being the people from whom the most co-operation is necessary to obtain access to data, they tend, at the same time, to be those who see the investigation as hostile to their interests.

### The Adversary Approach

Before leaving the topic of evaluative disputes, it is necessary to mention one approach to evaluation that sees dispute as an integral part of the evaluation process. Using an analogy between organisational decision-making and the process of legal judgement, the approach is sometimes known as the 'Adversary Model'.

For Levine<sup>17</sup> the need for an adversary approach seems to arise from a belief that experimental methods are inappropriate to the naturalistic settings in which most evaluation research is carried out. The most appropriate research instrument to deal with the complex social world, he says, is the equally complex mind of the human observer. This is highly sensitive, but is open to error which must be controlled. Non-quantitative research should be able to be open to procedures which enable estimates to be made of the confidence to be put in the conclusions.

To do this involves:

1. Appointing an Adversary to field and clinical studies whose job is to develop a rebuttal for all evidence gathered and introduced in support of a proposition
2. Appointing a Reviewer, to whom the Adversary can refer, and who "follows the primary investigator around, as much as is feasible, and who can provide an independent view of events".<sup>18</sup>
3. Since the subjective reactions of the investigator are part of the evidence, error must be discounted by becoming familiar with "the characteristic emotions, distortions, fantasies, defenses and values of the investigator."<sup>19</sup> This is to be done by a supervisory person or the group of investigators as a whole.
4. There is a need to develop a compilation of conditions of adequate inference and rebuttal for different classes of evidence. Levine also refers to the non-quantative school surveys, which were carried out in the 1920's in the USA, as fruitful sources of insight. These are the types of study which many evaluators regard as exemplary of the bad-old-unscientific-days.

Kourlisky<sup>20</sup> sees the adversary approach more as a way of presenting evidence for decision-making rather than as a model for research. She regards the approach as relevant when decision-makers do not want merely to rubber-stamp the recommendations of experts, and thinks it particularly appropriate when policy decisions involving large amounts of money are at stake.

The decision process is thought of as centring on a policy question, such as 'Should Programme X be continued?' Using the adversary approach, two Evaluators, the Affirmative and Negative, each present a written case to the decision-maker. The Negative Evaluator does not have to produce a counter-proposal, but simply an argument against the proposal. They then meet with the decision-maker. He cross-examines them; they cross-examine each other. The decision-maker tries to acquire an overall understanding of the main arguments upon which he bases his decision.

Kourlisky says that the approach rests on the notion that "Decision-making situations exist in which debate, dialogue, and discussion serve to improve the soundness of the judgemental process."<sup>21</sup>

Levine's prescription for the conduct of research studies is somewhat vague and is sometimes unrealistic, as in the case of point 3. But the idea of systematically identifying bias and of submitting evidence to the test of rebuttal seems sound. Kourlisky's recommendations also seem reasonable, though the rule that the Negative Evaluator should not argue for a counter-proposal, but simply rebutt the main proposal, seems unhelpful. This restriction means that the decision-maker is not informed of alternative courses of action. In the case of a dichotomous choice, each Evaluator is arguing for one course of action. But in a multiple choice situation, the approach should allow consideration of as many alternatives as are relevant. However, the systematic presentation of arguments and evidence for and against important proposals seems clearly desirable.

Whatever the drawbacks of the adversary approach, it does have one particular advantage. By encouraging as many arguments and as much evidence as possible for and against a proposal to be put forward, it can perform the important task of 'clearing the air', and enable all the interested parties to obtain a reasonably comprehensive view of the problem and its possible solutions. In a sense it forces each party to make a case, and at least makes it possible for the parties to be aware of each others views.

REFERENCES

1. D. Barker, 'Favourable Motorway Meals Survey Hard to Swallow',  
The Guardian, 19 April 1974, p.7.
2. M.D. Shipman, The Limitations of Social Research, London 1972.  
See also D.L. Phillips, Knowledge From What? Theories and  
Methods in Social Research, Chicago 1971, and Abandoning Method,  
London 1973.
3. Shipman, pp.150-51.
4. Ibid., p.155.
5. Ibid., p.92.
6. National Referendeum Campaign, Why You Should Vote No. and,  
Why You Should Vote Yes. London 1974.
7. Meehan, p.107.
8. Stake (1967), p.526.
9. Wiseman and Pidgeon, p.90.
10. Ibid., p.59.
11. Coombs, p.7.
12. Provus, p.180.
13. Weiss, p.18.
14. M. Mieux, 'Resolving Value Conflicts', in Metcalf.

15. Mieux reports that "the attempt at conflict resolution may result in an increased understanding of and respect for the other person, a kind of 'getting below the surface'." Mieux, p.164.
16. C. Wright-Mills, The Sociological Imagination, Harmondsworth, Middlesex 1970, pp.88-89.
17. M. Levine, 'Scientific Method and the Adversary Model: Some Preliminary Suggestions', Evaluation Comment, 4:2 (1973).
18. Ibid., p.2.
19. Ibid., p.2.
20. M. Kourlisky, 'An Adversary Model for Educational Evaluation', Evaluation Comment, 4:2 (1973).
21. Ibid., p.5.



## CHAPTER SEVEN

### THE EVALUATION OF OPEN UNIVERSITY COUNSELLING:

#### I THE BACKGROUND

In this and the following chapter, an account will be given of a particular evaluation research project which was undertaken at the Open University over the period 1970-74. The purpose of these chapters is to provide an empirical account of a specific attempt at evaluation research, which can be examined in the light of the previous analysis of the evaluation process. Some of the problems encountered by the project will be examined, and it is hoped that it will help to meet the demand for case-studies of research voiced by such writers as Clark.<sup>1</sup>

The particular project to be reviewed is one which began as an evaluation of the Open University's counselling services. The writer's involvement was as a Research Assistant working on the project over the period January 1971 to January 1974. The account has been constructed largely from documentary sources together with the recollections of the writer and others.

To begin with, it is necessary to say something further about the nature of the Open University, and the origins and development of its counselling services.

The Open University is a unique institution within the British system of education. Its basic function is to provide degree-level courses for adult students who may have no educational qualifications, and the courses are intended to enable students to pursue the bulk of their studies at home. According to Crispin,<sup>2</sup> the first public airing of the idea for an Open University (then called 'The University of the Air') was in a speech made by Harold Wilson in Glasgow in 1963. In 1966, the Government of the day issued a White Paper on the subject which recommended the setting up of a Planning Committee. This was established in 1967 under the then Vice Chancellor of the University of Aston, Sir Peter Venables. It issued a report in 1969 which outlined a comprehensive, but tentative, plan for the University, "leaving the elaboration of the detailed blueprint to the Council and Senate of the University".<sup>3</sup> This report was immediately accepted by the Government, and a Charter was granted on May 30th 1969.

The first teaching year began in January 1971, amid uncertainties as to the University's future after the Election victory of the Conservatives the previous year. The initial intake consisted of 25,000 students.

The University aims to provide generalist rather than specialist degrees, and a BA or BA with honours is awarded on the completion of a given number of courses. Each course has a credit rating, and students qualify for the award of a degree by accumulating a specific number of credits.

The courses are constructed by course teams composed mainly of members of the University's Faculties. There are six Faculties of Arts, Social Science, Science, Mathematics, Technology and Educational Studies, each headed by a Dean. As well as Faculty members, course teams may also include representatives of the BBC, members of the University's Institute of Educational Technology, and outside consultants. The course teams work to produce a number of Course Units which go together to make a course. Each Unit consists of a correspondence package (usually a printed booklet), and a television and radio programme. Students may also be provided with such items as gramophone records, tapes, home experiments, self-administered comprehension tests, and tutor- and computer-marked assignments. The broadcast components of the courses are transmitted by the BBC, and the other materials are sent to the students' homes through the post. Many courses also have a provision for the students' attendance at a short residential summer school, where intensive face-to-face tuition and other activities can be provided.

The Open University's headquarters is situated at Milton Keynes in Buckinghamshire, and there are thirteen Regional Offices located in major cities. Each Regional Office is the administrative centre for a Region. In England, the boundaries of the Regions are close to those of the Regional Advisory Councils for Further Education, whilst Scotland, Wales and Northern Ireland are Regions in themselves. Each Region is headed by a Regional Director who is responsible to the Director of Regional Tutorial Services at Milton Keynes.

Within each Region, there are a number of study centres, about 260 in all. Originally these were envisaged as TV viewing centres, but they now serve a number of purposes. They normally consist of a number of rooms hired by the University from a local educational establishment, and are equipped in most cases with TV and radio receivers and a library of course materials. Some also have replay devices for video and audio tapes, and computer terminals. Each study centre is open most evenings of the week, and they are manned by members of the University's body of part-time staff. The University employs about 5000 of these, and on each evening of the week a number of them are available at study centres for consultation and work with students.

There are two basic kinds of part-time staff. One, called a Course Tutor, is responsible for marking and commenting on students' written assignments, which are sent to the tutor through the post, and for providing occasional face-to-face tuition at study centres. Students may also write to or phone the Course Tutor for advice. The second part-time role is that of Counsellor. Briefly, his function is to act as a continuing personal link between the student and the University, and as a general educational adviser.

Part-time staff are normally attached to a Region, and are supervised by a number of full-time Staff Tutors and Senior Counsellors who are based at the Regional Offices. There are upwards of 150 Staff Tutors and Senior Counsellors, and as well as working with part-time staff in the Regions, they are responsible for their selection, briefing and training, and may also be involved in the production of course materials as members of course teams.

In addition to the six Faculties mentioned above, the University also has an Institute of Educational Technology which has parity of academic status with the Faculties. It was the Institute (IET) which became directly involved in the research work concerned with the evaluation of counselling.

The Institute of Educational Technology came into being in April 1970, almost a year after the University had received its Charter. The Planning Committee had foreseen the need for such a unit, which it referred to in its report as an operational research unit. Its proposed role included the study of new methods of teaching and communication, and experimental work, particularly in relation to the learning process.

During 1969, a firm of educational consultants (Instructional Systems Associates<sup>4</sup>) assisted the University in its operational planning, and they proposed the formation of an Applied Educational Sciences Unit to advise continuously on the design of components of the instructional system and to evaluate the system's effectiveness. This Unit began work in January 1970.

In the summer of the preceeding year, a Preparatory Courses Research Unit had been established. The Unit studied the workings of courses provided by the National Extension College and the BBC, which were similar to those which the University was intending to offer. When the Applied Educational Sciences Unit got under way, it was proposed that it should be linked with the Preparatory Courses Research Unit under the new title of the Institute of Educational Technology. This rearrangement was formalised in April 1970.

The work of the Institute can be divided into two broad areas. The first involves the continuous provision of advice to course teams on the design of educational materials and related matters such as evaluation and assessment. To this end, each Faculty is allocated a small number of Course Team Representatives from IET, who work closely with course teams. In 1971 there were about nine Course Team Representatives, currently about sixteen. The second area is that of research, including areas such as media research, textual communications, assessment, the structure of knowledge, student backgrounds and progress and tuition and counselling. The 'student based' research is largely carried out by the Institute's Survey Research Department, which is led by some of the original members of the old Preparatory Courses Research Unit. Those involved in the other research areas include Research Officers and Assistants, Lecturers, Senior Lecturers and consultants. Much of this research work is intended to assist the University's acting and planning.

The writer carried out research into the University's tuition and counselling services, as provided by the Regional part-time staff, as a member of the staff of IET.

#### The Background of Opinion

In order to understand the subsequent story of the attempts to evaluate the Open University's counselling service, it is necessary to describe the way in which this service came into being.<sup>5</sup>

The first mention of the counselling service is to be found in the 1969 Report of the Planning Committee. As has been noted, the Planning Committee only provided a blueprint for the University, leaving much of the detail to be worked out after it had been officially created. The specification of the local services (as opposed to the centrally produced and distributed course materials) was therefore rather vague. Mention was made of Local Viewing Centres (which became study centres), and of part-time tutors teaching through correspondence. The Committee also regarded as of "particular importance" the development of a counselling service, which would "help to reduce to a minimum the number of students who embark upon courses only to find that they cannot continue with them."<sup>6</sup> No specific mention was made of a person to be called a 'counsellor', and the Committee could be interpreted as meaning that counselling would be essentially something the student received before beginning studies with the University. After this, it would be a part-time tutor's job to "deal with the students' written work, and to guide and counsel them about their studies and progress."<sup>7</sup>

Late in 1968, the Planning Committee appointed six Directors of Studies. One of these was the Director of Studies, Local Centres and Tutorial Services<sup>8</sup>, whose job was to organise the corps of part-time tutors, the counselling service, the Local Viewing Centres and Summer Schools.<sup>1</sup> So far, the Planning Committee had said that there should be part-time tutors teaching through correspondence and a counselling service, and that "as far as possible there should be face-to-face as well as postal contact between the student and tutor."<sup>9</sup> The

Director had to translate this blueprint into an operational plan, and then into a working reality.

In April 1969, the Director sent a paper on the regional tutorial system to the Planning Committee. In it he said that face-to-face tuition would be a relatively rare commodity. The students' main human contact with the University would be a counsellor, who would direct students to tutors for face-to-face tuition as necessary. At a Senate meeting in June 1969, the counsellor was regarded as the main human contact with the students, with specialist tutors only rarely available to meet students at local centres. At the same time, Instructional Systems Associates were providing analytical reports on various aspects of the University's operations. They were not asked to examine the local tuition and counselling services, but their view was that at-a-distance teaching could work well enough to minimise the need for local support.

In September 1969, the Director sent a paper to the Vice Chancellor's Committee on the subject of personal tuition,<sup>10</sup> clarifying his earlier paper to the Planning Committee. He noted that some people may have interpreted his earlier paper as advocating "the substitution for the specialist tutor of a non-specialist (and therefore it was implied inferior) tutor in the form of the academic counsellor."<sup>11</sup> He declared that this was not his intention, and that "personal tuition was regarded as an essential feature of the system",<sup>12</sup> the problem being how to provide it within the constraints of finance and circumstances. He suggested that the variety of tutors a student would necessarily encounter as he changed courses might lead to



"a disturbing sense of discontinuity."<sup>13</sup> Moreover, "adult education experience and common-sense ..... show that many of the educational needs of the part-time student do not demand the attention of an academic specialist, and that consequently it would be a wasteful use of his services to employ him to cater for these needs."<sup>14</sup>

The idea of an 'academic counsellor' was therefore put forward. His primary function would be of "a general educational rather than academic specialist nature."<sup>15</sup> However, the problem of specialist academic tuition remained. It was noted that face-to-face tuition offered the greatest opportunity of individualised attention among the University's teaching media, but it was also suggested that "personal tuition to be truly individuated should only be provided for revealed or discerned personal needs."<sup>16</sup> Furthermore, "to attempt to provide the same amount of personal tuition for each student runs counter not only to the realities of the student situation and tutor resources, but also the very function of such tuition in the system as a whole."<sup>17</sup> Personal tuition might thus be provided for groups of students with common needs which could not be met by the counsellor. The counsellor would inform the Regional Office of the existence of such groups, and a specialist tutor would be provided if resources allowed.

The Vice Chancellor's Committee considered the Director's paper and noted the main conclusions, namely that provision could be made for face-to-face tuition on a limited basis within the financial estimates, and that such a system would enable such tuition to be provided for those students needing it. It also noted that:

"The case for face-to-face tuition should be established by the course team only after detailed consideration of the material to be taught, and the particular needs of OU students. Falling back on face-to-face tuition so early might suggest that the course team was taking the easy way out. On the other hand, the present lack of expertise in course design, which would obviate the need for face-to-face tuition, was a serious problem."<sup>18</sup>

It is interesting to note that at this time, face-to-face tuition was regarded as not to be the norm, as available only to students who needed it and on a limited basis, and as being potentially dispensable once expertise in course design had been acquired.

At the same meeting, it was proposed that a Tuition and Counselling Project Working Group be established. Its remit was to consider the provision of tuition and counselling within the constraints of the estimates. The Group met for the first time on October 16th 1969, and met on twenty-two occasions between then and November 1970, when it became reconstituted as the Tutorial Board. The Group originally consisted of the Director of Studies, Local Centres and Tutorial Services, a member of what was later to become the Institute of Educational Technology, representatives from each of the course teams then in existence, and a Regional Director.

At the Group's second meeting, the Faculty members (i.e. those from the course teams) emphasised the desirability of face-to-face tuition, but were warned that excessive reliance on it would disadvantage some students. This was because it was thought likely

that some students would not be within easy reach of study centres, or be able to go to them if they were. It was suggested that face-to-face contact between student and tutor was desirable fortnightly for students taking Maths and Science courses, and monthly for those taking Arts and Social Science. The notes of the meeting recorded that:

"It had originally been envisaged that students would have their counsellor as their major point of contact with the University. The consideration now being given to the provision of face-to-face tuition seemed to some members to relieve counsellors of some of their duties, and raised the possibility that a smaller counselling staff would suffice. Other members of the Group were strongly opposed to any reduction in the counselling service."<sup>19</sup>

The members from Science and Maths Faculties thought that "the financial provision for the counselling service could be raided to support face-to-face tuition. Some members expressed strong reservations at this suggestion."<sup>20</sup>

Thus began a long argument over the relative merits of tuition and counselling which culminated, as will be shown, in the setting up of a Review Group on Tuition and Counselling in 1973.

The third meeting of the Group reflected the same argument:

"The Group was reminded that, although the course teams had expressed strong support for an element of face-to-face tuition, it was open to members to recommend that no such provision be made. The

University was already spending a considerable amount integrating correspondence teaching with TV and radio broadcasts and a whole spectrum of educational technology. If existing correspondence colleges could teach by correspondence alone, the degree of sophistication being built into OU courses should be sufficient without any face-to-face tuition at all."<sup>21</sup>

At its fourth meeting, the Group considered a paper from the Regional Director of the East Anglian Region. This paper, entitled 'Student Counsellors: Their Function and Workload' outlined a number of possible functions for the counsellor. It then suggested that it would be impractical for specialist tutors to execute these functions, and gave a number of reasons why adult students needed counselling. Reference was made to the high failure rate of part-time students, and the last point was that:

"The Regional staff, on the best advice available from those who are expert in the field of the education of the adult part-time student on high level courses, have consistently argued for the importance to the student of having access to an educator who is regularly available throughout the student's course. This has been the conception which has been discussed with University Extra-Mural Boards, WEA and Local Education Authorities when we have been asking for their assistance in recent months. So far, these bodies have not questioned the conception; indeed, they helped to advise on its creation from their own predominant experience of adult students."<sup>22</sup>

On December 17th, the Senate considered a draft report prepared by the Group.<sup>23</sup> This outlined the level of tuition and counselling services to be provided for students, and the principles upon which the recommendations were based. These were that students should be able to see a counsellor at least once a fortnight at a study centre, with other forms of contact for non-attenders; tutors should be provided to mark students' written assignments; and a remedial tuition service should be provided for students who were experiencing particular difficulties in following the academic material. These services were regarded as essential, but it was although thought highly desirable that students should have regular face-to-face tutorials with subject specialists. It was recognised, however, that such tutorials could not be made available to, nor be used by, all students, and that they were to be regarded as an extra.

Counsellors were to be responsible for groups of about twenty students, ten of whom he could meet each week. Students would in many cases have the same counsellor for the whole of their time at the University. Counsellors would be part-timers having the educational qualifications and experience necessary to guide and help students. Where the counsellor's academic qualifications coincided with the discipline of an OU course, students taking such a course would, if possible, be allocated to him.

The specific functions of counsellors envisaged at this time are shown in Table 2.

Table 2: Envisaged Functions of Counsellors in 1969

1. Personal continuing relationship with the student - contact and encouragement
2. To help students with study problems. This need not imply specialist knowledge
3. To deal with the personal and social needs of the student, including domestic and personal problems
4. Help, where necessary, in running study centres
5. To create conditions in which students may meet informally
6. To ensure conditions whereby resources of study centres - radio, TV, library, computer terminals and other aids - are available to students
7. To create conditions for informal student group discussions
8. To guide students on remedial help available in the University
9. To keep personal records of students and ensure these records are used as instructed by the Regional Director
10. To make contact with students unable to attend study centres
11. To prevent drop-outs and follow up non- or irregular attenders
12. To advise students on vocational and further education opportunities
13. To advise students about further OU courses

Details were also given of the proposed functions of the tutors who would mark assignments, called correspondence tutors, and of the Regional supervisors who would oversee the work of the part-time staff. The face-to-face tutors, called class tutors, were envisaged as providing:

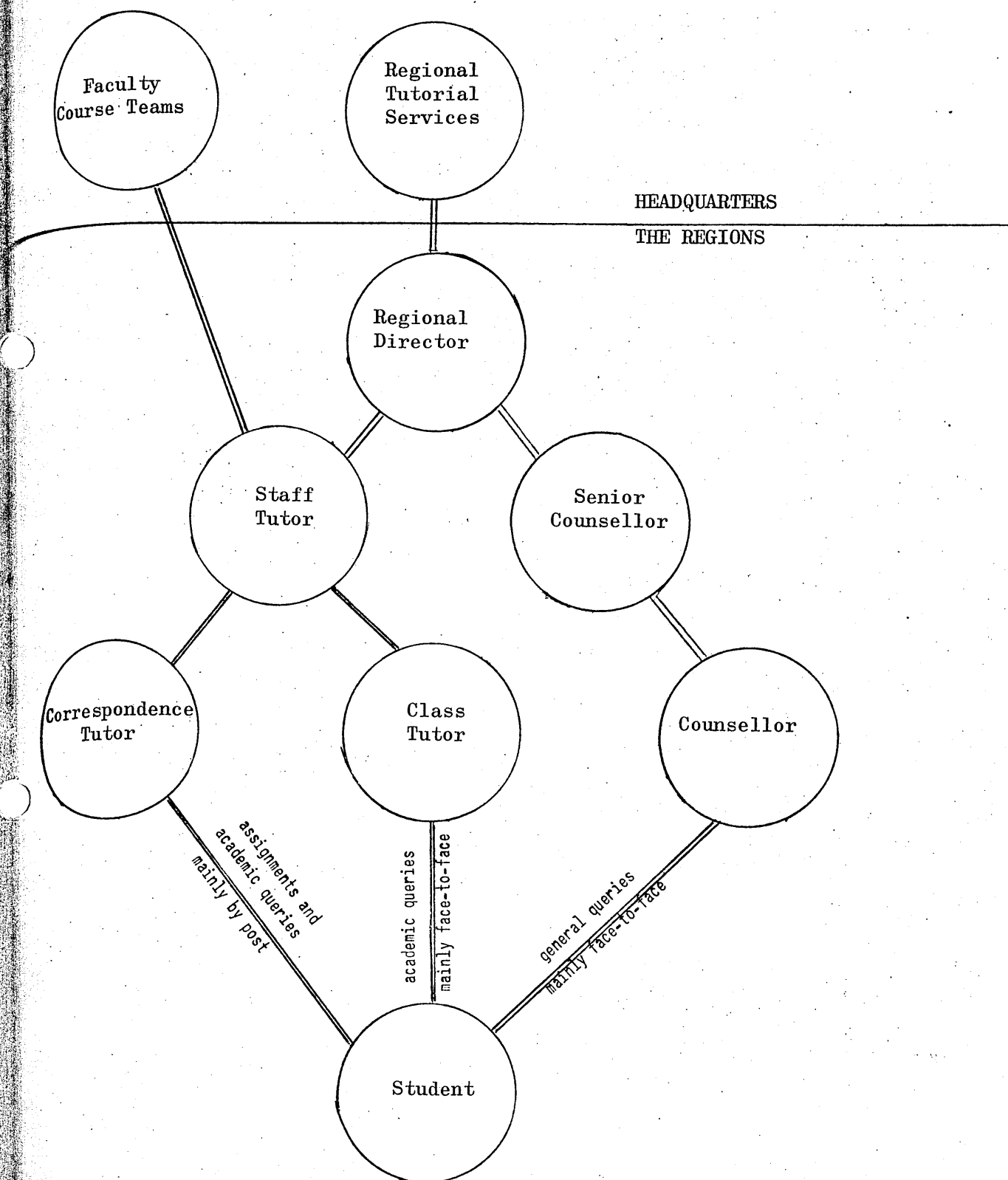
1. Face-to-face tuition to groups of students
2. Individual encouragement and guidance of students in groups
3. Liaison with counsellors, including written reports on students' progress and difficulties
4. Feedback to the course teams on student reaction to course material

It was noted that individual members of part-time staff might take on more than one of the three roles of counsellor, correspondence tutor and class tutor.

The Senate agreed the scheme put forward by the Group, but pointed out that it might not be possible to establish the services at the level proposed, for organisational and financial reasons. It therefore referred the proposals back to the Group for priorities to be established. The Senate's view was that priority should be given to the provision of a counselling service and to correspondence tuition. As Hawkrige<sup>24</sup> has noted, face-to-face tuition was being played down at this time.

In March 1970, the Senate received a further report from the Group. The scheme proposed was basically similar to that put forward the previous December, although the specific duties of the class tutor and the correspondence tutor were somewhat fluid.

Figure 11: The 1971 Tuition and Counselling System\*



\*See also Appendix III.



The Senate agreed the proposals, which then formed the basis for the construction of the tuition and counselling services for 1971. Briefly, then, these consisted of:

A counsellor, whom the student could see once a fortnight and who would deal with study and administrative problems

A correspondence tutor, responsible for grading and commenting on students' written work, and for replying by letter to queries about the courses

A class tutor, who would hold tutorials at the study centre based on the written course material or students' queries

The Group now turned its attention to a host of additional details, on the basis of a general agreement on the basic structure of the tuition and counselling services. These included details of the payment system for part-time staff, and the drafting of 'further particulars' for both part-time staff and the full-time Staff Tutors and Senior Counsellors.

At its fifteenth meeting, held on April 27th 1970, the Group noted that there was a need to formulate a detailed budget for tuition and counselling for consideration by the Council and Senate of the University in June. At its sixteenth meeting, the following estimates were presented:

Correspondence tutors	£564,000
Class tutors	£ 91,400
Counsellors	<u>£326,000</u>
Total for part-time staff	<u>£981,400</u>

These costs represented a substantial proportion of the University's overall budget.

The Dean of the Science Faculty commented that the proposed expenditure on counselling was unjustifiably high in relation to the total. In his view, the funds for counselling would be more profitably spent on increasing the provision of face-to-face tuition. He did not wish, however, to reopen the discussion about the structure of the local services, since this had already been agreed by the University. But he did consider that research should be conducted into the effectiveness of the counselling service. The Group as a whole agreed that it was in favour of any research that would assist the University in coming to informed decisions about the allocation of tuition and counselling resources. It was therefore proposed that the Director of the Institute of Educational Technology (which had, it will be remembered, an institutional research type brief) should be asked to attend the next meeting to discuss ways in which a research programme for 1971 could be organised.

The events following on this request that research be conducted into the effectiveness of the counselling service will be described in the next chapter. For the moment we can note some of the general features about the debate over the form which the University's local tuition and counselling services should take.

One of the important factors which influenced the whole planning process of the Open University was the almost total novelty of the enterprise. The core of the idea for the Open University was teaching at-a-distance; the ability to reach students in their own homes, in all parts of the country, primarily via the media of broadcasting and the written word. In the rather traditional environment of the British higher education system, the idea was totally unconventional, and there were few precedents upon which the planners could draw. At the same time, the University had to be designed in a very short period. It is a remarkable fact that only two years after the Planning Committee had issued its tentative blueprint, the first OU students were receiving their packages of course materials.

The novelty of the institution and the pressures of time inevitably necessitated some fairly robust planning, and provided considerable scope for debate about how the University was to be organised. The basic idea of using broadcasts and correspondence materials to teach students at-a-distance, with tutors marking assignments and returning them to students through the post, was fairly easily established. But what of the more traditional forms of face-to-face tuition, and the general educational adviser and personal counsellor? These were contentious issues, both with regard to the desirability of providing them in a fundamentally distance system, and with regard to the practicality of providing them. Would it be possible to find enough tutors in the right

place and with the right discipline to ensure face-to-face tutorials for most students? Would students face enough personal and study difficulties to warrant the provision of a personal counsellor? Would, in fact, the knowledge that there were teachers available to students locally prove a disincentive to course teams to produce courses that communicated first time? Should all available resources be ploughed into the central production process with no local contact at all? All these were questions of the moment for which there were no concrete answers.

The Director of Studies, Local Centres and Tutorial Services, believed in the need for a general educational adviser to students, with some face-to-face tuition if circumstances allowed it. Faculties, on the other hand, were more emphatic about the high desirability of face-to-face tuition, and some of them regarded it as more important than counselling. Others were concerned lest the local services became a Frankenstein's monster, devouring resources which could be used to so improve course materials as to make local support unnecessary, and perhaps diminishing the University's potential as a distance teaching institution.

According to the Vice Chancellor,<sup>25</sup> the provision of face-to-face tuition for students was not thought to be a practical proposition, partly because it was not known whether there would be enough specialist tutors living near enough to students to enable them to meet. The Planning Committee's remit for face-to-face contact between tutor and student remained, however, and as the Tuition and Counselling

Project Working Group considered the form of local services, the Faculties made clear their desire for face-to-face tuition.

The need for some accessible human adviser to students, available locally, was accepted. Since the possibility of providing local tutorial support was doubted, at least on any significant scale, the counsellor was regarded as the key role. When the provision of face-to-face tuition began to seem a realistic proposition, this seemed to some to undermine the case for having a counsellor.

The upshot of the argument was a compromise, and the system established for 1971 embodied both face-to-face tuition from class tutors and general guidance from counsellors. As it turned out, this was to be an uneasy compromise.

#### Changes in the System

The very early experience of the system in 1971 resulted in some immediate proposals for change. The Tuition and Counselling Project Working Group was reconstituted as the Tutorial Board in January 1971 at the beginning of the University's first teaching year.

A sub-committee was set up to prepare proposals for tuition and counselling for 1972. These had to be approved by the Senate in March 1971 to enable the necessary arrangements to be made for the following year.

Two problems seem to have emerged, although there was some disagreement over the evidence of their existence. One was that correspondence and class tutors felt isolated from each other.

Policy at that time was that no student should have the same person

as both their class and correspondence tutor, since it was felt to be advantageous for the student to be able to receive tuition from someone who was not also his assessor. The class tutor worked mainly at the study centre, whilst the correspondence tutor worked mainly from home, so the opportunities for direct contact were limited. A second problem was that counsellors felt the need for copies of the course materials their students were working on, which as non-specialist, general advisers they did not at that time receive.

The proposals for 1972 therefore advocated a change from three roles - class tutor, correspondence tutor, counsellor - to two. Each student would now have:

A course tutor, who would be a subject specialist and who would both mark students' written work and give face-to-face tutorials. The advantages of combining assessor and tutor were thus seen to outweigh those of separating them.

A counsellor, who would carry on his general adviser role as before, but who would have a 'defined tutorial role' in respect of students taking a Foundation course (the lowest level course provided by the University). The counsellor would be provided with course materials, but was expected to use these to help students with problems of study technique, rather than subject matter problems. For post-Foundation level students, he would have no 'defined tutorial role'.

Interestingly, the sub-committee which prepared these proposals reported that in some of its discussions it had been "hampered by semantic difficulties."<sup>26</sup> Such difficulties were to bedevil discussions on tuition and counselling.

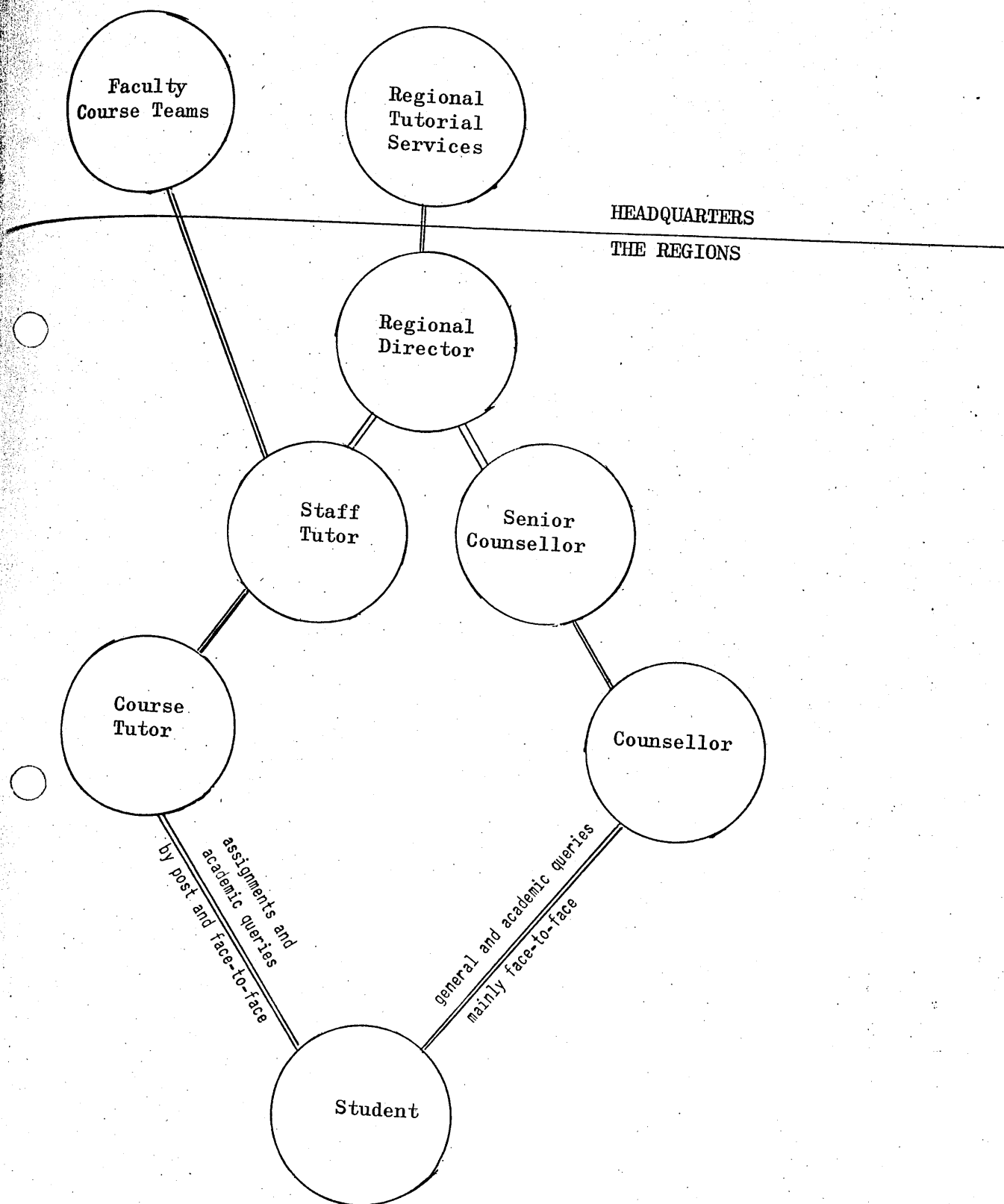
After a very full discussion, the Tutorial Board agreed the new scheme at its meeting of March 9th 1971. There were, however, dissenting voices. Social Science Faculty Board said:

"The Board felt that there were problems about the proposal to extend the role of the counsellor to include some form of tuition; after all, counsellors were recruited because they had certain characteristics which did not necessarily guarantee their suitability for handling the tutorial side. Moreover, it was clearly impossible to come to a definite view about the counsellor's role after only seven essentially atypical weeks of operation. The Faculty Board recommends that a decision be held over pending a more reliable assessment of the success or otherwise of the counselling function."<sup>27</sup>

There was certainly a dilemma here, between waiting longer for more evidence of how the system was working and risking being unable to implement changes until 1973. In any event, the changed system was established in 1972, and continued more or less unchanged through 1973 and 1974. Further changes are, however, planned for 1976. The diagrams on pages 264 and 272 show the structure of the tuition and counselling arrangements before and after the 1972 change.

After 1972, changes in the course distribution of students led to further consideration being given to the need for changes in the tuition and counselling arrangements. In 1971, all students had been taking Foundation courses, and the large numbers taking

Figure 12: The Tuition and Counselling System 1972-75\*



\*See also Appendix III



any one course made it possible to arrange meetings between tutors and reasonably sized groups of students. As the number of courses available increased year by year, the number of students on any one course decreased, and in some cases there were as few as two hundred students registered for one course. Because such students were scattered all over the country, it became more difficult to arrange for face-to-face meetings on a viable basis. To overcome this, provision was made for flexible distribution of the hours of face-to-face tuition, so that some Regions would provide a small number of day schools, others the normal evening tutorials, and others a mixture.

These difficulties began to provide a case for a more substantial funding of post-Foundation, as opposed to Foundation, level tuition. The case was given added weight by the discovery that the drop-out rate (the proportion of students failing to complete courses) among post-Foundation students was higher than had been expected. Moreover, some of the Faculties which had always been sceptical about the need for counsellors brought increasing pressure to bear on the counselling budget. Crudely stated, their case was that counselling was either ineffective or unnecessary, and that what the students needed was more tuition. As has been said, counsellors were not expected to give tutorial type help to post-Foundation students. Therefore it was the counselling budget that was attacked. The necessity for this arose partly from the financial stringency imposed on the University by Britain's severe economic crisis, which made the prospect of any major increase

in the overall budget of the University unlikely. There was even a possibility of cuts. Internal redistribution therefore seemed the best prospect.

Thus, three years after his initial criticism of the counselling budget in the Tuition and Counselling Project Working Group, the Dean of the Science Faculty wrote in May 1973:

"We would like to be able to assume that, by 1976, counselling will either have been made far more effective than it appears to have been so far, or be drastically curtailed and the resources thereby saved be used to increase tutorial help (especially at-a-distance tuition) to the students most in need of it."<sup>28</sup>

The pressure for change reached a peak in the Spring of 1973. The Tutorial Board had already agreed to transfer £30,000 from counselling to tuition at post-Foundation level, and as the discussions about the policy for 1974 tuition and counselling took place, the Arts Faculty made two proposals.

Firstly, it wanted the counsellors' 'defined tutorial role' dropped for students taking the Arts Foundation course. That is, it did not want counsellors to give subject matter tuition to the Arts Foundation course students. Its argument was that "this would have the merit of reducing the amount of face-to-face tuition, which had been felt by some to be excessive, and of simplifying the counsellor's function by removing ambiguity."<sup>29</sup> It also wanted a review of the counselling service. At the same time, the Science Faculty proposed

an increase in the number of students allocated to a counsellor, the financial savings being diverted to tuition.

The most significant of these proposals, and the one that was agreed, was the proposal to set up a Review Group. This Group was instituted in April 1973, with the title of the Review Committee on Tuition and Counselling. Its remit was to devise alternative models of tuition and counselling, within the assumptions that there would still be study centres, a counselling activity, and the same level of correspondence tuition as in 1973. Any new proposals had to cost no more than the system already in operation.

It is not necessary to go into the details of the Review Committee's work, but simply to note a few important points which will be commented on later.

The Review Committee's recommendations represented something of a compromise between those who wanted to see counselling severely pruned and those who wanted it to continue basically unchanged. The Committee's Interim Report (September 1973) outlined three models of tuition and counselling. Of the three, the Committee favoured a 'combined model' of tuition and counselling.

This involved replacing the course tutor and counsellor with a new role called the academic supervisor. The academic supervisor would perform most of the course tutor and counsellor functions for both Foundation and post-Foundation level students. It was thought that there would be less emphasis on counselling for post-Foundation

level students, on the assumption that they would be more experienced and would thus need less general assistance. However, to assist the academic supervisor there would be a Study Centre Counsellor who would be responsible for monitoring the progress of individual students and dealing with particular problems referred to him by the academic supervisor. Unlike the existing counsellor, however, who might have between twenty and thirty students to deal with, the Study Centre Counsellor would be responsible for up to 250 or so students. The Committee described this model as "a radical departure from the present system."<sup>30</sup>

Initially, any new model for tuition and counselling was intended to be implemented in 1975. When the Review Committee's Interim Report was circulated, there was a demand for a lengthier period of consultation over the proposals, which were then rescheduled for implementation in 1976. In its Final Report, the Committee noted that none of its proposals had commanded overwhelming support when first presented. It had, therefore, on the advice of the Tutorial Board, now produced a model which it unanimously endorsed.

This model embodied two main provisions. Foundation course students would have a combined course tutor and counsellor. One person would perform both functions. Post-Foundation level students would have separate course tutors and counsellors. The principle of 'continuity of counselling', ie. the belief that each student should have a link with one member of staff throughout his career at the University, would be preserved by students retaining their course tutor-counsellor as their counsellor when they moved beyond Foundation

level. In the model proposed in the Interim Report, this principle had been thought by some to be endangered, as it was not believed that a counsellor with a 'case load' of several hundred students could develop any real personal relationship with individuals. On the other hand, the new model also helped to satisfy the demand for more tuition by eliminating the situation in which Foundation level students received a double dose of tuition via their course tutor and the counsellor's 'defined tutorial role'. The latter, although intended to mean that course tutors would give subject matter advice and counsellors general study advice, was believed in practice to mean that counsellors were giving basically tutorial help. Hence the Arts Faculty's view that the amount of personal tuition at Foundation level was excessive. The combination of roles at Foundation level was intended to overcome this problem, and released funds to increase tutorial provision for post-Foundation level students. It was re-emphasised, however, that the main instrument of the University's teaching system was the at-a-distance approach, especially correspondence tuition.

The Committee also noted that there was "a shortage of precise information regarding the way in which the tuition and counselling system works in practice",<sup>31</sup> and proposed that a new communications system should be devised. It concluded by saying that "these arrangements should be reviewed after a period of 3 - 5 years", and that "during that period, there should be an evaluation of the revised system."<sup>32</sup>

Comments

The debates over the arrangements for tuition and counselling at the Open University, took place in an essentially political context. It is not simply that there was disagreement about the best system, but rather that the subject was overlain with emotional connotations of the kind one might associate with the symbolism of ideological conflict. Thus it was not uncommon to hear the argument couched in terms of an 'attack' by the Faculties on the Regions, or of the Centre on the Periphery. The supporters of counselling were sometimes referred to by their opponents as patronising, tender-minded student molly-coddlers, proffering shoulders that no one wanted to cry on, whilst in turn, the tutorial lobby was sometimes regarded as a group of inhumane, unsympathetic academics, more interested in shoring up the defects of their courses with local tutorial support than with the needs and problems of individual students. Some believed that there was a distinct and important role for counsellors, and thus regarded the addition of a 'defined tutorial role' for 1972 as 'selling counselling down the river'. Others thought the whole idea of counselling was a mistake.

This atmosphere is well reflected in the thesis by Crispin referred to above. Writing as a Staff Tutor at the Open University, his thesis centred on the role of the part-time tutor. He writes:

"We shall recall the early decisions which, if anything, demoted face-to-face tuition in favour effectively of counselling, and trace the stages of its re-emergence as an important feature of the learning system."<sup>33</sup>

He then goes on to talk about tracing the "revival" of face-to-face tuition. Commenting on the revised system for 1972, he says:

"With respect to the tutor, this reorganisation could be interpreted as a further step towards his reinstatement, and a recognition of the importance of the face-to-face teaching approach."<sup>34</sup>

He later asks why the Faculties had not originally "campaigned" for face-to-face tuition.

Certainly the argument over tuition and counselling was one which raised many passions, and the importance of the controversy was remarked on by the Director of Studies, Local Centres and Tutorial Services, in a memo to the chairman of the Review Committee in October 1973:

"The need to unite the University, and especially the members of Regional Academic Staff most directly concerned with tuition and counselling, is, in my opinion, no less important now than the need to radically improve the present system. I believe the model we propose can do both. I hope it will be considered by the Review Group in the spirit in which it is put forward as a way towards consensus in an area that has for too long been the subject of debilitating controversy which, if it persists, may

well damage the University's service to its students and place the well-being of at least the most vulnerable in jeopardy."<sup>35</sup>

The form of the local tuition and counselling services thus became a particular focus of argument for the following reasons.

Firstly, there was the fact that the Planning Committee's original blueprint for the University left the precise structure of the local services unclear. This meant that at the beginning, there was little in the way of an authoritative specification which could be used as an uncontroversial starting point. Similarly, when the detailed planning began, the lack of precedents for the University to follow meant that there was insufficient evidence to direct the debate forcefully in one particular direction. There was also a great pressure to complete the planning quickly, which prevented some of the kinds of investigations and inquiries which might have resolved some of the factual issues. Thus there was plenty of room for disagreement about what would be the best system.

There could not, of course, have been an argument without there having been proponents of conflicting points of view. Since the argument was essentially about the level of resources which should be allocated to face-to-face tuition as against counselling, it was natural that the disputants should be drawn from those whose interests and responsibilities were most closely associated with each. The organisational structure of the University was conducive to polarisation along these lines.



The Director of Studies, Local Centres and Tutorial Services, had been one of the foremost advocates of counselling. The Regional Tutorial Services division of the University, of which he was the head, helped to execute tuition and counselling policy and provided administrative support to the Regions. The Senior Counsellors, based at the Regional Offices, were, with one or two exceptions, staunch supporters of counselling. Some of the Faculties, on the other hand, wanted to promote face-to-face tuition at the expense of counselling, and the Regional Staff Tutors tended to take the same view. There thus arose what might be loosely described as a Faculty-Regional split, the former supporting tuition and the latter counselling. It is perhaps worthwhile noting that the Institute of Educational Technology's view was that, in principle, students needed neither face-to-face tuition nor counselling, and their continuance reduced the incentive to produce self-sufficient teaching materials, and for that matter self-sufficient independent learners. This faith was demonstrated in one of their own course proposals, which made no bid for face-to-face tuition (the provision of counselling being beyond the control of course teams).

Another contributing factor, possibly the major factor, in the dispute was that of money. It will be remembered that it was the cost of counselling that provoked the Dean of Science to request research into the effectiveness of the service. In a new enterprise whose justification was, at least in part, its potential cost-effectiveness in providing degree-level education, there was a

considerable amount of cost consciousness. On top of this, the advent of a Conservative Government and Britain's perennial economic crisis, did not bode well for any significant increase in the University's overall budget. Thus internal redistribution was the alternative. The tuition and counselling services were a particularly likely candidate for redistributions because they formed a major part of the University's direct student costs. These costs varied with the total student population, unlike those arising from activities such as course production which remained constant irrespective of student numbers, and the tuition and counselling system was therefore likely to take a bigger and bigger share of the cake as time went on and student numbers increased. There was therefore considerable concern about whether these costly services were adequate. Some felt that it was anomalous to spend so much on local support in what was intended to be a distance teaching system.

With this background established, we can now go on to examine the attempts to evaluate Open University counselling under the auspices of the Institute of Educational Technology's Tuition and Counselling Research Project.

REFERENCES

1. Clark, p.146.
2. A. Crispin, 'Part-time Teaching Within the Open University System with Particular Reference to the Role of the Part-time Tutor', MA (Ed.) dissertation, University of London, (1973), unpublished.
3. Department of Education and Science, The Open University: Report of the Planning Committee, para. 5.
4. Two members of this firm became full-time members of staff of the Institute of Educational Technology.
5. A chronological summary of the main events associated with the development of the tuition and counselling system is given in Appendix II.
6. Department of Education and Science, The Open University: Report of the Planning Committee, para. 51.
7. Ibid., para. 72.
8. The title was later changed to Director of Studies, Regional Tutorial Services.
9. Department of Education and Science, The Open University: Report of the Planning Committee, para. 72.
10. Director of Studies, Regional Tutorial Services, 'An Approach to the Question of Personal Tuition', memorandum to the Vice-Chancellor's Committee, 15 September 1969.
11. Ibid., para. 1.

12. Ibid., para. 1.
13. Ibid., para. 2.
14. Ibid., para. 2.
15. Ibid., para. 3.
16. Ibid., para. 4.
17. Ibid., para. 4.
18. Vice-Chancellor's Committee, Minute 76.0.2, 25 September 1969.
19. Tuition and Counselling Project Working Group (T&CPWG),  
Notes, 6 November 1969.
20. Ibid.
21. T&CPWG, Notes, 13 November 1969.
22. D. Grugeon, 'Student Counsellors: Their Function and Workload',  
undated, p.4.
23. T&CPWG, Report, VC/12/1, 8 December 1969.
24. D.G. Hawkrige, 'The Role of the Course Tutor in the  
Open University', paper presented at the Open University  
Conference on Correspondence Tuition, June 1973.
25. Open University, The Early Development of the Open University:  
Report of the Vice-Chancellor, January 1969 - December 1970,  
Bletchley, Bucks. 1972, p.66.
26. Tutorial Board paper, TB/3/1, 9 March 1971.

27. Social Science Faculty Board, Minutes, 24 February 1971.
28. Science Faculty, Science Foundation Course Remake Proposal,  
May 1973.
29. Tutorial Board, Minutes, TB/17M, 5 March 1973, 21.9.
30. The proposals were, of course, much more detailed.  
Only an outline has been presented here.
31. Review Committee on Tuition and Counselling, 'Interim Report',  
September 1973, para. 2.3.1.
32. Review Committee on Tuition and Counselling, 'Final Report',  
February 1974.
33. Crispin, p.5. My emphasis.
34. Ibid., p.20. My emphasis.
35. Director of Studies, Regional Tutorial Services, memorandum to  
the Chairman of the Review Group on Tuition and  
Counselling, 12 October 1973. The model proposed  
was one created by the Director and his staff.

## CHAPTER EIGHT

### THE EVALUATION OF OPEN UNIVERSITY COUNSELLING:

#### II TUITION AND COUNSELLING RESEARCH 1970-73

In May 1970 the Dean of the Science Faculty asked that research be conducted into the effectiveness of the counselling service. As we have seen, this request was passed to the Institute of Educational Technology. In this chapter we will describe the events which followed this request, and some of the activities involved in the subsequent research programme. Although many topics were examined by the researchers concerned we will concentrate mainly on those relevant to counselling. Since the research was conducted on a yearly time-scale, the chapter will be divided into four sections each covering a single year.

#### 1970: The Beginnings of the Project

Nineteen-seventy was the run-up year to the Open University's first teaching year, and was thus an extremely busy one. As we have seen, one of the areas of concern at this time was the organisation of the local tuition and counselling services. By May, the Tuition and Counselling Project Working Group had considered the financial implications of its proposals, and research into the effectiveness of the counselling service had been requested. The Group agreed that it would welcome research

which would help the University to make informed decisions about the allocation of resources to tuition and counselling, and asked the Director of the newly formed Institute of Educational Technology to attend its next meeting to discuss the possibilities.

To see why the Institute of Educational Technology was regarded as the natural basket into which this particular egg should be placed, it is necessary to consider the way in which the Institute saw its role at this time. We have seen how the Institute came into being via the merger of the Applied Educational Science Unit and the Preparatory Courses Research Unit. The Director of the new Institute arrived in January 1970, and like the rest of the University, the Institute was involved in sorting out its roles and functions.

It is probably fair to say that the Institute saw its role as something of a crusading one at this time. In a way, the Open University must have been seen by educational technologists as a dream come true, and the Institute saw its task as that of promoting the systematic approach to teaching among staff who were less experienced in innovatory methods. At the same time, it saw itself as the guardian of the institution's wellbeing which it thought would be best served by ensuring that hard evidence was collected to inform future decisions. A draft press release, prepared in November 1970, read:

"The Institute staff believe that whatever the University produces must be subject to constant testing and revision. They want decisions about

changes in the courses to be taken on the basis of hard data rather than on subjective opinion."

And the actual release read:

"With such facilities available there can be less dependence on intuition to the improvement of teaching methods, although leaps of intuition will still be required. Our diagnostic analyses should give evidence of the effectiveness of different course segments, and possibly on the effectiveness of different media. Similarly, it should be possible to compare the success of different groups of students and the effectiveness of tutors.

Our aim is to design an instructional system that can improve itself. The OU is the first University in the world to base its improvement strategies so much on objective data."<sup>1</sup>

The Institute saw itself as having an important role to play in many areas of the University. For instance, a set of particulars for posts in the Institute, prepared in September 1970 read:

"At present, the Institute's members are working chiefly in the areas of resolution of objectives, preparation of tests, and developmental testing of course units. Work has begun on a practicum for the use of Open University course teams. Attention will be paid shortly to problems of integrating the media in the Open University instructional system. Individual staff members are becoming deeply involved in fundamental planning of many other areas of the system as it evolves."<sup>2</sup>



Given its wide research interests, the Institute was the obvious locus for the kind of research the Tuition and Counselling Project Working Group envisaged.

On June 8, members of the Institute met to discuss "possible crises in the University's budget." It was noted that owing to shortage of funds, some areas of the University's operations were in danger of being cut. Some Faculties had attacked the level of expenditure on counselling, "but these Faculties tended to have a subject-matter oriented approach, which stressed the need for highly qualified tutors, while possibly neglecting the human problems of the adult student at-a-distance." On the other hand, class tuition had now come to prominence at the expense of counselling. The meeting concluded that "unless the Institute, or some other body within the University, undertook some investigations on the effectiveness of personnel, informed decisions in the future would be impossible."<sup>3</sup>

Having decided that the Institute was in favour of the kind of research requested by the Tuition and Counselling Project Working Group, the Director of the Institute sent a memorandum to the Group. This began by saying that "the Institute's staff are absolutely in agreement with [the] suggestion that research should be conducted into the effectiveness of the counselling service", and that the Institute hoped to "turn its attention to the counselling service very soon." A number of general observations about counselling were made, with suggestions for a diary study and the collation of informal feedback with the help of Regional staff. It was stated

that the Institute would be responsible for designing the research materials, for collecting, collating and analysing the data, and for preparing reports for University committees. It was also stated that "IET concern will lie in the area of general evaluation of the counselling service and all aspects that make it effective or ineffective, rather than in the designation of individuals as effective or ineffective."<sup>4</sup> Finally, an offer was made to prepare a research proposal.

The Tuition and Counselling Project Working Group considered the memorandum on July 9, but in the meanwhile it had prepared an interim report on its work for the Senate. This included a section entitled 'Research and Evaluation', and indicated that a request for research had been made to the IET. However, the brief had now been widened to include all the main components of the local services - "The Group attaches particular importance to the evaluation of the effectiveness of the three main elements of the tutorial service-counselling, class tuition and correspondence tuition."<sup>5</sup>

At its July meeting, the Group accepted the broad proposals put forward by the Director of IET and asked for further details. It was concerned that any research should proceed "in close cooperation with Regional staff." It also agreed to investigate the possibility of finance for the research from the University's budget for local services.

Meanwhile, plans were going ahead in IET for the creation of a Tuition and Counselling Research Unit, whose chief objective

was to be "to evaluate the tuition and counselling services through monitoring the activities of correspondence and class tutors and of counsellors."<sup>6</sup> At this time it was thought that the Unit would be staffed by short-contract consultants and Research Assistants.

At a meeting in August, the Tuition and Counselling Project Working Group considered the Institute's evaluation plans (see Appendix IV). The Group accepted the plans, subject to the availability from the budget for local services of the £14000 estimated as necessary for their execution. The Group also reiterated its view that "any research programme should be an evaluation of the system only, and in no way a monitoring of individuals within the system."<sup>7</sup>

By September, the possibility of finance for the research was still in doubt. The Director of IET submitted new estimates to the Group giving a reduced figure of £10000. He pointed out that the basic question was whether or not to spend £10000 evaluating a programme costing £1.4M.

On October 13, the Senate (the University's supreme decision-making body) received another report from the Tuition and Counselling Project Working Group. In this the Group strongly recommended that approval be given for the proposed expenditure for the research. Senate agreed to this and referred the matter to the University's Planning Board, a body responsible for long-term planning and the allocation of resources. However, by October 30, the Director of IET noted that no provision seemed yet to have been made in the estimates for the research. He now proposed that the

expenditure should be £10000 in 1971 and £5000 in 1972 and 1973. He also pointed out that "to indicate how seriously the Institute views the need for the evaluation of the tuition and counselling system, .... the Institute has made an offer of one of its Research Assistantships in this area to an individual this week,"<sup>8</sup> the writer being the individual concerned.

The Tuition and Counselling Project Working Group agreed to put the matter before the Planning Board at once, and it was subsequently approved.

Work now began on formulating detailed research plans. At a meeting of the Institute in December, it was announced that money to finance the Tuition and Counselling Research Unit had been obtained, and that the Unit consisted of the Director of the Institute, three other Institute staff and a Regional Director. The only full-time member of the Unit would be the Research Assistant. Preliminary plans were discussed on the basis of the objectives stated in the August proposal. The Institute approved these, adding that "additional emphasis be laid upon the impact of the tuition and counselling system on the learners."<sup>9</sup>

The idea of research Units in the Institute did not emerge in practice, since some members felt that this would encourage staff to have overly narrow concerns. Those involved in the tuition and counselling research later became known as the Tuition and Counselling Research Planning Group, and this was headed by the Director of the Institute.

Thus by the end of 1970 the initial request for the evaluation of the effectiveness of the counselling service had been widened into an evaluation of all of the main elements of the local services; the evaluation was to be of the system only and not in any way a monitoring of individuals; plans had been prepared in the Institute of Educational Technology, funds obtained and a small staff allocated to the project.

### The Project in 1971

The first phase of the project involved sending questionnaires to tuition and counselling staff in order to determine their knowledge of their roles and the problems they anticipated. The first recipients of these questionnaires were the Senior Counsellors and Staff Tutors, who supervised the part-time tutors and counsellors in the Regions. Regional Directors and Deans of Faculties were circulated with details of the objectives of the research in order to keep them informed.

At about the same time a proposal emerged that pairs of Staff Tutors and Senior Counsellors be invited to visit a study centre in a neighbouring Region to add to their own experience and to record 'successful' tuition and counselling practices. The Regional Director on the Planning Group was particularly keen on this, and it perhaps reflected the concern that Regional staff should be involved in the research work. This Regional Director later became an Assistant Director of Studies working with the Director of Studies, Local Centres and Tutorial Services at the

University's headquarters, and thus provided a most important link with the central tuition and counselling administration.

Unfortunately this particular project turned out to have unexpected repercussions.

In February a conference of Senior Counsellors was held to discuss various aspects of tuition and counselling. It was attended by most Senior Counsellors as well as some Staff Tutors, Faculty members from HQ, and Regional and Assistant Regional Directors. At this conference the Director of IET presented a paper outlining the proposals for tuition and counselling research,<sup>10</sup> and requesting cooperation for the proposed exchange visits to study centres. The paper is a particularly important one since it was here that the Director outlined his conception of evaluation. It is therefore necessary to give a brief summary of it.

The paper began by reminding everyone that under the extreme time pressures which had been imposed on the planning of the University, many decisions about its organisation had had to be taken on the basis of inadequate information. "Some of you", said the Director, "may have an appreciation of the combination of sheer hard planning, lengthy committee discussions, political manoeuvring, and general muddling through and compromise that has stilled the crises to some extent and allowed the Open University to get off the ground." He was, however, alarmed that "we have neglected to build into the system the proper ways of improving it .... we are still faced to operate a system without being able to find out systematically whether the system is operating well.

Finding out what is working well is what I call evaluation."<sup>11</sup>

The role of the Institute was then outlined, and a request made for cooperation in the planning and execution of the evaluation of tuition and counselling, which IET could not hope to do alone. Evaluation was not just concerned with providing a better basis for decision-making, but was also a matter of "political wisdom". "The University, like all Universities, is a political arena. Persuasive powerful figures win what they want unless one can marshal evidence against them."<sup>12</sup> Hard objective evidence was needed, and this evidence needed to be of high quality. He continued:

"Let me end my apologia by suggesting too that the evaluation of the tuition and counselling system is in your own interests. Many people in the University have already noted a debate going on between the regional organisation and the central administration. I have heard it said in the regions that decisions are being made at the centre without any concern for the actual conditions in which the regions operate. Whether or not this is true, there is no doubt that a proper evaluation of the tuition and counselling system will increase the chances of well-informed decisions at the centre."<sup>13</sup>

The Director then described the objectives of the research, and indicated some of the problems in achieving them, particularly the problem of getting results in time for decisions. The first part of the project, involving the questionnaires sent to part-time staff mentioned above, was already under way. It was acknowledged

that some senior staff might have thought these burdensome during the initial crisis period of the University (which was worsened by a postal strike at the start of the year). The administration of some of the questionnaires had also been delayed because of difficulties in securing a suitable sampling frame from the University's Data Processing division.

The proposal for visits to study centres was then mentioned, and it was intended to discuss the details at a later point in the conference. Perhaps regional sub-groups could be set up to advise on the evaluation of tuition and counselling. Perhaps other projects beyond those planned could also be done. Finally the problems of data-collection, retrieval, analysis and interpretation were summarised. The Director concluded:

"It may be several years before we can hope to have neatly proscribed projects with experimental and control groups and so on. Perhaps these will always be out of place at the Open University."<sup>14</sup>

The immediate reaction to the paper was rather lukewarm, and one or two people expressed some hostility to the idea of evaluation. On the whole the reaction might best be described as neutral. Nevertheless, the proposals for study centre visits were discussed later, and seemed generally acceptable to those present. Unfortunately, most of the Staff Tutors who were likely to be involved were not present at the conference.

The basic format for the study centre visits was that pairs of Staff Tutors and Senior Counsellors would visit a study



centre in a neighbouring Region on a number of evenings, and write a report on their observations. It was hoped that this would include a general description of the study centre and its activities, information obtained from part-time staff and students, and accounts of 'successful practices' of part-time staff. The resulting material was then envisaged as being useful for the briefing and training of those involved with the local services.

The Regional Director on the Planning Group had now become Assistant Director of Studies, as mentioned above, and he took on the job of arranging the pairings for the visits. The IET staff prepared a short manual for use during the visits, giving a format for reporting and some advice on interviewing drawn from a standard source. This was sent to all the participants in the project.

By the middle of April it was becoming clear that the study was running into difficulty. Staff Tutors in the Science Faculty refused to cooperate with the visits, since they felt that there was some implication that they were unfit to assess matters in their own Region. It is certainly difficult to see how they arrived at this view. They also felt that they were being ordered around without consultation.

Some of the Regional Directors were also unhappy. One wrote to say that the research materials should have been sent to him for distribution to his staff, rather than direct to them. Moreover, on behalf of his staff he took "considerable exception to the fact that it was thought necessary to spell out some quite

elementary points for the benefit of academic staff who have now conducted a number of visits to study centres since January this years." He warned IET to "tread carefully"<sup>15</sup> when approaching Regional staff.

Another Regional Director wrote in the following terms:

"I have literally been assailed by the strongest possible objections from Staff Tutors and Senior Counsellors to the paper ..... quite unprecedented anger created by the manner in which the entire project has been presented so far .... resentment ... disillusionment ....."<sup>16</sup>

Some of the Faculties also had complaints. The Science Faculty wrote to say that it had "strong objections" to the study centre visits, that it had not been consulted about the use of its staff, and that it wanted to be consulted in future. The Dean of the Science Faculty said he was "very worried" about the project, and concluded:

"All in all a sadly bungled business I fear. We can try to pick up the wreckage and start again, but it will not be easy."<sup>17</sup>

The Dean of the Arts Faculty also said that this Staff Tutors were reacting unfavourably and were offended.

In response to this reaction, the Director of IET circulated a clarificatory memorandum. He explained that no one was under an obligation to participate in the study, and that if people wanted to write up their existing experience that would be fine.

The research 'manual' had been meant for guidance rather than instruction and simply outlined standard research practice. As a result, a number of visits to study centres did take place, and the resultant reports were edited and distributed to Regions as briefing material.

Although the hostile reaction was a minority one, the researchers were nevertheless surprised by its vehemence. Indeed it was rather unexpected since the actual focus of the study seemed fairly innocuous. Talk of "unprecedented anger" and picking up "the wreckage" seemed quite out of proportion to the 'offence'. Still, the researchers recognised the need to avoid such situations in the future, however 'irrational' the reactions appeared, and thus steps were taken to try to prevent it happening again. Until now, the Planning Group had been composed largely of IET staff with a solitary Regional Director. It was now decided to expand the group so as to include a number of Senior Counsellors and Staff Tutors. As there were about two hundred of these all told, only a small proportion of them could be included, particularly since it was felt to be important to keep the Planning Group fairly small. Thus about half a dozen joined the group.

A meeting of the Group took place on May 25. The problems that had arisen over the visits were discussed, and a number of reasons were put forward to explain them. These are summarised in Table 3. The lesson that emerged from this was that obtaining data was likely to be a much more difficult exercise than had been thought at first, and that henceforth consultation and liaison

Table 3: Explanations of Difficulties Associated with  
the Study Centre Visits Project

1. A failure to communicate the aims of the research adequately
2. Inadequate cooperation between IET and the central Regional administration
3. The feeling among Staff Tutors of being used as IET research assistants. Senior Counsellors had a formal obligation to work on 'institutional' research projects written into their contracts, but Staff Tutors did not
4. Objectives were not spelled out clearly enough, whereas the method was spelled out too clearly, so it was hard for participants to see the point of the research
5. Staff Tutors and Senior Counsellors were already overloaded and saw the visits as an unnecessary frill
6. There was a conflict between the objectives of enabling senior Regional Staff to visit another Region and of getting specific data for training purposes. Regional Directors were nervous of being compared unfavourably with another Region by virtue of the visitors having visited a 'dud' study centre, or having gone on an 'off' night. There were worries about upsetting the activities at a centre
7. People were sensitive because they were struggling just to get the University going

would be vitally important. At the February conference, the Director of IET had said that "in this whole project we are trying to measure the size and tell the shape of an extremely jellyish jelly-fish."<sup>18</sup> It now seemed as if this jelly-fish was of the stinging variety.

In view of these problems, the writer produced a paper which focussed on some of the human relations problems involved in evaluation research, drawing on Havelock's<sup>19</sup> work on innovation in education. It attempted to suggest some ways of structuring the relationship between the research group and the 'clients', and suggested the need to decide what might be involved in evaluating tuition and counselling, since this was the formal brief. However, the paper did not lead to any change of strategy.

The rest of the year was spent working on the variety of separate projects which has been thrown up as the year progressed. A list of 'successful counselling practices' was derived from the reports of the visitors to study centres, and circulated to Regions.<sup>20</sup> The questionnaire to part-time staff was also analysed and the results similarly circulated. One of its main conclusions was that the tuition and counselling system seemed to be operating as planned. It was left to the Faculties to decide if there was a need to change recruitment policy to part-time posts in the light of the data provided on the backgrounds of part-time staff.<sup>21</sup> Various feedback forms were considered for analysis, and reports were prepared on two of those used by counsellors.<sup>22</sup> A limited evaluation of the University's summer schools was conducted, and

an analysis made of the relation between performance ratings of summer school tutors and background variables.<sup>23</sup> An outside consultant also did some work on the University's assessment system.

In the autumn of 1971, the Director prepared an Annual Report on the tuition and counselling research.<sup>24</sup> In that year no less than fifteen 'phases' of research had emerged. Of these, eight had been or were nearing completion, one had been transferred to other IET researchers, and the rest were in the early stages. In view of the fact that there had been only one full-time researcher working on the project, it was planned that a Research Officer be appointed for the following year. Moreover, less than £2000 was spent from the budget of £10000.

Of the difficulties reported, one was particularly significant. Throughout the year various requests had been received by the Planning Group, but there had been:

"Difficulties in obtaining a coherent brief from the University's committees and in developing a coherent plan of action; for example, agreement about summer school evaluation was not reached until too late a date for much action to be taken."<sup>25</sup>

The Report emphasised the point in relation to plans for the following year:

"Phases (and priorities) cannot be finally agreed until there are reasonably detailed statements of aims, justifications, requirements and constraints.

The requests made by the Tutorial Board at its last meeting are not sufficiently explicit, for example."<sup>26</sup>

Unfortunately these requirements were not to be met.

To summarise, 1971 saw the execution of a wide variety of projects, some of which were concerned with counselling. Some difficulties were experienced with gaining cooperation from the various parties to the research, and steps were taken to try to remedy some of the main problems as the researchers saw them. The budget was heavily underspent, but it was planned to increase the manpower on the project for the following year.

#### The Project in 1972

The Tuition and Counselling Research Planning Group, consisting of IET members, Staff Tutors, Senior Counsellors and the Assistant Director of Studies, met on January 3. The possibility of constituting the group as a formal University committee was considered. So far the group had had a semi-formal status, and it was thought that by making it a formal committee it would gain in status as a legitimate originator of research. The Director of IET's view, however, was that the group would probably operate more effectively on an informal basis, unencumbered by the trappings of the University's administration. It was agreed that written notes of the group's meetings should be kept in future and that these would be circulated to the group's members. It was also agreed that the Tutorial Board, from the which the group received its brief and to whom it reported,

should clarify its problems to a greater extent before submitting them to the group.

The plans for 1972 covered thirteen projects, some of these being repeats of those carried out the previous year or completions of unfinished work. Some of these had been specifically requested by other University bodies, and some were suggested by the group. The proposed projects are shown in Table 4. Some of these projects were to be carried out by the Director of IET, the Research

○ Assistant or Research Officer; others would be carried out by other IET members or consultants. Tentative priorities were attached to the projects (mainly High) and the Tutorial Board, to which the proposals were submitted for comment and approval, was asked to clarify the objectives for those projects which it had suggested. Hopeful of a clear response, the proposals were sent to the Tutorial Board.

The Tutorial Board was a large and heavily burdened committee, and its response was unenlightening. It generally  
○ agreed all the proposals, and tried to raise the priority of non-High projects to High, thus negating the purpose of assigning priorities. The Board was requested to set up a small group to formulate objectives for the unclear projects, but the Board did not do this. Nor did it offer clarification itself.

As a result, the Planning Group went on to formulate objectives for some of the projects, and submitted these to the next meeting of the Board. When the Planning Group next met,



Table 4: Proposed Projects for Tuition and Counselling  
Research for 1972

1. Continuing survey of the backgrounds of part-time staff
2. Studies related to the development of briefing and training materials for part-time staff
3. Further monitoring of the University's feedback system
4. Studies of the relation of assignment cut-off data to student performance and retention in the system
5. Analysis of a set of record forms kept by Senior Counsellors
6. Completion of analysis of two types of record form kept by counsellors in 1971
7. Studies of students' assignments
8. Summer school evaluation
9. Study of the need for preparatory courses
10. Study of students not using tuition and counselling services
11. Study of the use of playback and viewing facilities in study centres
12. Study of the use of audio-cassettes by Technology Faculty students
13. Collection and analysis of feedback from tutors involved in trial of two science courses.

it noted that:

"The Tutorial Board had acknowledged the proposals and not commented adversely or otherwise. It was to be assumed, therefore, that they had been accepted."<sup>27</sup>

In the same month (March) the project's new Research Officer arrived, and he was given the task of organising the summer school evaluation study.

The rest of the year was spent working on the projects generally agreed by the Tutorial Board. One other important event was the arrival in April of a second Assistant Director of Studies. The Director of Studies, Local Centres and Tutorial Services, now had two Assistant Directors. The first now became generally concerned with the tuition side of the local services, and the newcomer was to be largely responsible for counselling. She joined her colleague as a member of the Planning Group.

It is not possible here to describe the progress of all the studies conducted by the group, and since we are mainly concerned with counselling we will consider only those projects which had some bearing on this area. We will not go into the details of the data here, as this will be considered as a whole later.

Briefly summarised, the studies having some relevance to counselling were as follows:

The survey of part-time staff: a short questionnaire was sent to all members of part-time staff, asking them about their

employment background and some personal details. A very high response of 90% was obtained, and the resulting tabulations were distributed to Faculties, Regions and various University committees.<sup>28</sup> The tables included figures for counsellors. An analysis was also made of drop-out and turnover among part-time staff.<sup>29</sup>

Analysis of a small number of forms used by counsellors to report problems to Senior Counsellors: the results were inconclusive since many counsellors were believed to report the most important problems by telephone. The analysis of a larger sample was not thought to be worthwhile.

Analysis of a form used by counsellors to summarise contacts and problems raised by students: this analysis was made by a member of the Institute's Survey Research Department, using a sample of forms completed by counsellors in 1971. In his report, the researcher expressed a number of reservations about the validity of the data.

Analysis of forms used by counsellors to report on their work in general at the study centre: these had been completed by counsellors late in 1971, and the information they contained was coded by Senior Counsellors using a coding scheme prepared by the writer. A summary report was circulated to Regions and to University committees.

Study of students not using tuition and counselling services: one of the Senior Counsellors on the Planning Group conducted a small study of the relation between attendance at a study centre,

the distance a student lived from the study centre, and the award of a course credit. No written report was produced.

This represents the main work carried out by the IET researchers in 1972 that was related to counselling. No Annual Report was written that year, although the Research Officer did produce a summary of the studies underway in September 1972.<sup>30</sup>

### The Project in 1973

Discussions on plans for 1973 began in the autumn of 1972.

The Director of IET proposed the projects listed in Table 5.

Day-school evaluation and Regionally based research were also mentioned as possibilities. These proposals were discussed at a meeting of the Planning Group on September 19 1972. At the same time, the Research Officer put forward a proposal for a Ryans-type teacher characteristics study of tutors.<sup>31</sup> The group seemed to endorse the proposals, save for the monitoring of some of the counsellors forms. These were declared by the new Assistant

Director of Studies to be mainly intended for Regional use, and would be analysed by non-IET personnel.

At about this time, the full time research workers on the project (the writer and the Research Officer) were beginning to develop a consensus between themselves that the project as a whole was not altogether satisfactory. Researchers are, of course, prone to fits of grumbling,<sup>32</sup> but there was a feeling that there were certain fundamental problems surrounding the project which refused to go away. In particular it was felt that the overall

Table 5: Proposed Projects for Tuition and Counselling Research for 1973

1. Studies of part-time staff and wastage among the part-time staff population
2. Monitoring of the feedback system involving examination of report and record forms
3. Studies of correspondence tuition
4. Summer school evaluation
5. Study of the need for preparatory courses
6. Continuation of the study by Senior Counsellors of a study centre
7. Study of playback and viewing facilities at study centres
8. Study of the use of audio-cassettes.

role of the project was unclear and that the rather short-term and piecemeal approach to research planning involved the researchers in some unconnected projects of limited value. The researchers thus met with the Director of IET, who it will be remembered was in overall charge of the project, to discuss these problems, before the meeting of the Planning Group which would finalise plans for the next year. The Research Officer emphasised his desire to do high quality research in a clearly defined area in terms of a study of tutor characteristics. It was agreed that in future a smaller number of studies would be pursued, and that a rough division of labour would be:

Research Officer: summer school evaluation  
study of tutor comments on  
students' scripts  
study of tutor characteristics

Research Assistant: analysis of backgrounds, drop-out  
and turnover among part-time staff  
studies related to counselling

With regard to studies of counselling, the writer was at a loss to suggest ones that were either not already being done (such as the analysis of forms) or which would definitely be likely to be useful as opposed to simply interesting. This was partly because the writer's view was that it was up to the 'clients' to originate the problems. If this did not happen, it seemed that they were likely to agree to almost any study on paper, but perhaps be less interested when it came to securing their cooperation in carrying the study through. The Director took a different view. He felt that it was up to the researchers to take a lead in suggesting projects,

and it was thus suggested that a study of counsellor characteristics might be undertaken in parallel with the study of tutor characteristics. Both studies would use Flanagan's 'critical incidents' approach.<sup>33</sup>

With this agreement established, the writer prepared a document for circulation to senior Regional staff, outlining four possible areas for research into counselling. The paper began as follows:

"It seems to me that crucial to any decision about how research resources should be allocated is a prior decision about the problems which the application of such resources is designed to solve. Research resources are scarce. They should therefore be allocated to problems which will give a high degree of pay-off relative to expenditure. This implies that we have a suitable analytical grasp of the situation that will enable us to assess in advance the impact of the information we produce as a result of our research efforts. We want to be able to feel reasonably confident that, if our results are valid, they will have some application to significant decisions.

The first aim of our research should therefore be to diagnose the range of problems which face us."<sup>34</sup>

Four Big Questions about counselling were then listed.

1. What constitute effective counselling? How is the performance of counsellors to be judged?
2. Are there any particular characteristics which distinguish more effective from less effective

counsellors? Are these characteristics which we can select for, or can they be acquired through briefing and training or both?

3. What tools, strategies and procedures can be devised which will help counsellors to perform their functions more adequately?
4. What do most counsellors spend most of their time doing and with what results?

I then outlined possible approaches to obtaining answers to these questions (see Appendix V).

The paper was circulated to all senior Regional staff, IET staff and the Planning Group, with a request for comment. Generally the response was a little disappointing, and mostly consisted of urgings to pursue one of the topics rather than others. Since there was no clear consensus for any one topic, most effort was put into the study of counsellor characteristics via the critical incidents approach.

The overall proposals were then submitted to the Tutorial Board (which referred them to one of its committees) and they were generally accepted.

At this time the first signs of the coming debate over counselling were emerging, and the opposing forces began to prepare themselves. In December 1972, the Assistant Director of Studies responsible for counselling issued a paper titled 'Some Developments of Counselling'.<sup>35</sup> This reviewed some of the changes which had been made in the counsellor's role over the previous



two years, and outlined some of the ways in which the role might be strengthened in the future. A second paper appeared in January 1973 called 'Summary Report on the Roles and Functions of Counselling and the Counselling Service'.<sup>36</sup> This was largely an expansion of the earlier paper.

By January 1973 the rumblings about counselling began to reach the researchers. On January 3, a discussion took place between the writer and the Director of IET concerning the proposals for counselling research. The Director thought that counselling would become a focus of debate when the proposals for the arrangements for local services were discussed by the Tutorial Board that coming March. As we have seen, he was correct in this. He thought that it was possible that counselling would be scrapped for 1974, and therefore wanted to delay the research into counselling until the picture was clearer.

On January 9, the head of the Institute's Survey Research Department wrote in response to the proposals for counselling research. She began:

"Your starting position is to my mind unexceptionable. The tragedy is that it was the starting position which we were at in 1970."<sup>37</sup>

In her view, the main task for counsellors was to stop students from dropping out of the University, the level of drop-out among students being of continuing concern to the University. The main task of the counselling research should be to "prove" that counsellors did this.

Meanwhile, the Tutorial Board's proposals for tuition and counselling policy for 1974 had arrived for consideration by the IET Board (the governing body of the Institute). One proposal was that £30000 should be transferred from the counselling budget to increase post-Foundation level support. As there was no extra money to be hand, and Foundation level tuition did not seem a good candidate for cuts, the obvious solution was to take money from counselling. The Institute's reaction to this proposal was to advocate a debit of £200000 from the counselling budget. This represented about 50% of the counselling budget as opposed to the 8% proposed by the Tutorial Board. The general rationale for this was as follows:

"It was agreed that the IET Board should press strongly the point of view that if more effort were given to improving the design of courses, then there would be less need for supporting services in the Regions."<sup>38</sup>

This reflected the belief of some members of the Institute that the counselling service was in some way inadequate or unnecessary. There was also the fact that the Institute was becoming grossly overloaded, underfinanced and understaffed. If money could be released for the improvement of course design this might mean some expansion of the Institute's resources.

In the same month a seminar was given by the writer and the Research Officer on the project, called 'Tuition and Counselling Research: What Should We Be Doing?'. This was later

written up as a paper by the Research Officer, and as it summarises many of the problems faced by the project, it is reproduced in full as Appendix VI. The idea for the seminar arose from the increasing awareness among the researchers that neither we, nor anyone else, really seemed to know what we should have been doing, and that non-researchers did not seem to understand the problems of conducting research in our particular context. It was thus an attempt to induce a greater understanding among the 'clients' of why we were doing what we were.

The paper began:

"It sounds like a good idea to have a project with a general responsibility for looking into the University's tuition and counselling services. There are obviously many questions in this area which it would at least be interesting, and in some cases extremely useful, to be able to answer. On the other hand, research costs money, takes time, tends to be intrusive and cannot guarantee to come up with answers. How much should we spend on research? Which questions should we tackle first? And what research methods should we use?".<sup>39</sup>

And among the conclusions were:

"..... we need to take rather more account in our discussions on research, of what specific questions we are trying to answer, what kinds of answer we are interested in, and whether we have methodology available to give that kind of answer."<sup>40</sup>

Unfortunately the seminar was rather poorly attended. Only about a dozen people came, including a Regional Director, one or two members of the Planning Group and some IET staff.

The seminar was heard sympathetically, but no solutions were offered. The paper was later discussed by the Planning Group where reaction was much the same. The notes of the May meeting record:

"The problems covered in the paper were acknowledged to be significant ..... It was felt that the 'clients' should try to formulate their problems adequately before 'commissioning' research into them. The group recognised the difficulties of determining research priorities, and of designing studies which would solve the problems that were brought to the group."<sup>41</sup>

These sentiments were almost identical with those expressed in the 1971 submission of research proposals to the Tutorial Board.

During the early part of the year some further discussions were held with the Director of the Institute concerning the counselling research for that year. He was concerned lest the research be made obsolete by events after it had half started, and since his view was that counselling had been misconceived from the start, he was concerned in case the research might strengthen the hands of those who wanted to retain it. It was therefore agreed that the work should be held over until after the March Tutorial Board meeting when the future of counselling would be discussed. It was at this meeting, as we saw in the last chapter, that the Arts Faculty asked for counselling not to execute their 'defined tutorial role' for Foundation level Arts students and

requested a review of the counselling service.

A few days after this meeting, a further discussion was held with the Director. He recounted the substance of the proceedings and said that there had been some complaints about the tuition and counselling research. These had come especially from the representatives of the Science and Technology Faculties. A set of reports produced by the project were thus to be sent to these critics and a summary of the data available on counselling produced.

A set of reports were sent to the Science Faculty representative, who replied as follows:

"I am certainly aware, and I'm sure others are as well, that IET had some studies of the counselling system and of counsellors in progress. However, again speaking for myself, I find it very embarrassing to be asked to comment, advise and reach a judgement on counselling and counsellors because hard information with which to work is so scarce or even absent altogether. The recent discussions arising from the Tutorial Provisions paper highlight the situation. We were asked to approve the spending of £400000 on counselling. The impression I have from students and staff tutors is that counselling is of minimum value and hardly used."<sup>42</sup>

The Science representative then cited the difficulty of obtaining information by reference to a memorandum which he had sent to the Director of Studies, Local Centres and Tutorial Services. He had asked for information on the number of contacts

between counsellors and Science students, the time taken with each, and the types of problems raised. The Director of Studies had replied by saying that there were no figures available at present.

The Science representative went on to make his Faculty's position clear:

"Now to add another twist to the problems in this area, the Course Teams are struggling to devise effective, interesting and novel ways of enabling students to learn. However, as we all know, resources are rationed and, naturally, hands are reaching out to transfer part of the money earmarked for counselling into other parts of the tuition budget. You can appreciate that it is an extra frustration, hardly to be borne, to find that transfer of money is held to be impossible for no reason that stands scrutiny."<sup>43</sup>

After the Tutorial Board meeting, it was clear that there were unlikely to be any major changes in the counselling arrangements for 1974. The study of counsellor characteristics thus went ahead and was finally written up in January 1974. The summary of data on counselling requested by the Director was completed.<sup>44</sup> This consisted of a collection of statistics drawn from previous reports, but excluding much of the student-based data collected by the Survey Research Department. The summary did not include any evaluative judgements. The Institute's Survey Research Department also produced a report based on data collected from students, and again the aim was to "present the data rather than interpret them".<sup>45</sup> These reports were sent to the Review Committee,

whose deliberations and recommendations we have already considered.

We are now almost at the end of the account of the Tuition and Counselling Research Project. Most of the 'routine' research projects were completed during the course of the year, bar the tutor characteristics study which had to be held over for another year because of pressure of other work.

○ At the end of 1973, the writer left the project to take on other work in the Institute. The project continued into 1974 much as before, although there had been some discussion about the possibility of transferring the work to other hands. There had always been some ambiguity about the proper locus for tuition and counselling research. The work of the Survey Research Department seemed to overlap with that of the project, as did that of the media researchers in the Institute. There was also the fact that the central Regional administration and Regional staff also had a direct interest in research into their own affairs. ○ There was thus not only a question of What Should We Be Doing?, but of Who Should Be Doing What?. The extremely burdensome duties of the Director of the Institute were added to by his role as director of the project, and although he had managed to take an active part in the research it was becoming more difficult for him to play an active role. In June of 1973 he thus suggested transferring the project to the Survey Research Department, but by July this no longer looked practical.

The money freed by the departure of the Research Assistant was used to fund two Senior Counsellors to conduct some research into student drop-out. The increasing concern of the central Regional administration to be able to conduct its 'own' research was evidenced by their appointing their own Research Assistant early the following year.

Our account of the Open University's Tuition and Counselling  
○search Project up to 1974 is now complete. It is now possible to examine some of the problematic aspects of the project, and in particular the problems of evaluating counselling. The account that has been given is obviously incomplete: four years cannot easily be compressed into a few pages. Nevertheless, it is to be hoped that enough has been said to lend substance to the remarks which follow.

○



REFERENCES

1. Draft and actual press release, issued 26 November 1970.
2. Institute of Educational Technology (IET), 'Specific Information on the Institute of Educational Technology', part of a set of further particulars for posts in the Institute, September 1970.
3. IET Board, Minutes, 8 June 1970.
4. Director of IET, Memorandum to the T&CPWG, TC/15/3, 16 June 1970.
5. T&CPWG, 'Report of the Tuition and Counselling Project Working Group to Senate', S/15/3, received by Senate 30 June 1970.
6. Director of IET, 'Research Plans for 1970-71', memorandum to IET staff, 16 July 1970.
7. T&CPWG, Minutes, 5 August 1970.
8. Director of IET, Memorandum to the T&CPWG, 30 October 1970.
9. IET Board, Minutes, 14 December 1970.
10. D.G. Hawkridge, 'Introduction to Some Aspects of Institute of Educational Technology Research', paper presented at an Open University Conference of Senior Counsellors, 23-25 February 1971.
11. Ibid., p.2. He was, of course, simplifying his views for the benefit of a non-specialist audience.
12. Ibid., p.3.
13. Ibid., p.4.

14. Ibid., p.11.
15. Regional Director, East Midlands Region, Letter to the Director of IET, 26 April 1971.
16. Regional Director, South Region, Letter to the Director of IET, 26 April 1971.
17. Dean of Science Faculty, Memorandum to the Assistant Director of Studies, Regional Tutorial Services, 3 May 1971.
18. Hawkridge, 'Introduction to Some Aspects of Institute of Educational Technology Research', p.10.
19. R.G. Havelock, A Guide to Innovation in Education, Ann Arbor, Michigan 1970.
20. IET, 'Successful Counselling Practices: Tuition and Counselling Research Project, Phase 2a)', September 1971. See Supplementary Material.
21. IET, 'Tuition and Counselling Research Project: Report on Phase 1b)', July 1971. See Supplementary Material.
22. A.B. Thomas, 'Final Report on the Analysis of Counsellors' Form C4', January 1972. See Supplementary Material.  
A. Woodley, 'An Analysis of a Sample of Completed C3 Forms', March 1972.
23. A.B. Thomas, 'Summer School Tutors 1971', November 1971.
24. IET, '1971 Annual Report to the Tutorial and Research Boards on the Tuition and Counselling Research Project', undated.

25. Ibid., p.3.
  26. Ibid., p.8.
  27. Tuition and Counselling Research Planning Group (T&CRPG),  
Notes, 15 March 1972.
  28. A.B. Thomas, 'Tuition and Counselling Research Phase 1a) 1972:  
Census of 1972 Part-time Staff', November 1972.  
See Supplementary Material.
  29. A.B. Thomas, 'Tuition and Counselling Research Phase 1b) 1972:  
An Examination of Drop-out, Wastage and Conservation of  
Part-time Tutors and Counsellors 1971-72', April 1972.  
See Supplementary Material.
  30. R.A. Northedge, 'Summary of Current Tuition and Counselling  
Research Projects', September 1972.
  31. See D.G. Ryans, Characteristics of Teachers, Their Description,  
Comparison and Appraisal, Washington D.C. 1960.
  32. See G. Boalt, The Sociology of Social Research, Illinois 1973,  
for an interesting consideration of the factors influencing  
the motivation of social researchers.
  33. See J.C. Flanagan, 'The Critical Incident Technique',  
Psychological Bulletin, 41 (1954).
  34. A.B. Thomas, 'Research Into Counsellors and Counselling in 1973',  
November 1972, p.1.
  35. E. Cronhelm, 'Some Developments of Counselling', December 1972.
-

36. E. Cronhelm, 'Summary Report on the Roles and Functions of Counsellors and the Counselling Service', TB/16/5, January 1973.
37. N.E. McIntosh, 'Comments on Alan Thomas' Notes on Research into Counsellors and Counselling 1973', undated.
38. IET Board, Minutes, 29 January 1973.
39. R.A. Northedge, 'Tuition and Counselling Research: What Should We Be Doing?', 19 January 1973, p.1.
40. Ibid., p.5.
41. T&CRPG, Notes, 10 May 1973.
42. Science Faculty, Memorandum to the Director of IET, 29 March 1973.
43. Ibid.,
44. A.B. Thomas, 'A Summary of Some Research Data on Counselling 1971-72', March 1973.
45. N.E. McIntosh and S. Blacklock, 'Students and Counselling', first draft, undated, p.3.

## CHAPTER NINE

### PROBLEMS OF EVALUATING OPEN UNIVERSITY COUNSELLING

In the preceding chapters we have traced the development of the Open University's local tuition and counselling services and the progress of the Institute of Educational Technology's Tuition and Counselling Research Project from 1970 - 73. Some problems which arose during the work of the project will now be examined, particularly with reference to the evaluation of counselling.

The first question which concerns us is this: in what respect can the project be regarded as problematic? There are a number of criteria which could be used to answer this, but the one which I shall use is whether or not the project answered the question which it was brought into being to answer. Thus if the project answered the question which it was originally intended to answer it can be regarded as unproblematic. If it did not answer this question, it can be regarded as problematic. It is suggested that the project did not manage to answer the question that prompted its creation, insofar as we restrict ourselves to the topic of counselling. As has been said, it would probably be widely acknowledged (without undue self-congratulation) that much useful work was carried out under the auspices of the project. Yet the topic of counselling was one which was of concern to the University throughout the period considered, and which ultimately led to

the creation of a Review Committee on the whole subject of the local services. The question of the 'effectiveness' of the counselling service was thus a live one throughout the period considered. It was, as it were, waiting to be answered.

It is probably not controversial to say that the project did not give an answer to this question, but it is perhaps necessary to try to establish the point. It would be no surprise that this happened if one of the tasks of the project had not been to evaluate the effectiveness of the counselling service; or if this aim ceased to be relevant once the project got under way. However, it seems that this was one of the original aims of the project, and that it continued to be one throughout the period concerned.

The initial request from the Tuition and Counselling Project Working Group referred to research into the effectiveness of the counselling service. This request was passed to the 'research group', i.e. the Institute of Educational Technology, and the Institute's staff were "absolutely in agreement with [the] suggestion that research should be conducted into the effectiveness of the counselling service."<sup>1</sup> There thus seems to have been agreement that this was a topic to be pursued. Moreover, it seems that there was agreement that this work was to be couched in terms of an evaluation. The Institute said that "IET concern will lie in the area of general evaluation of the counselling service";<sup>2</sup> the Tuition and Counselling Project Working Group also referred to "the evaluation of the effectiveness of the three main elements of the

tutorial service",<sup>3</sup> one of which, of course, was the counselling service. Thus the chief objective of the project for 1971 was "to evaluate the tuition and counselling services";<sup>4</sup> for 1972, "to conduct analytical and evaluative studies of the University's tuition and counselling system".<sup>5</sup>

Given that the project was intended by both the 'clients' and the researchers to determine, as one of its aims, the 'effectiveness' of the counselling service, it is now necessary to establish that it did not achieve this task. Having done this, we can consider why this happened.

The judgement that the project did not manage to evaluate the effectiveness of the counselling service is clearly dependent upon a notion of what would count as an evaluation. The position adopted so far has tried to view evaluation as more than just description, although descriptions are seen as very necessary to reasoned evaluations. In keeping with Scriven's view, it is proposed that there is no evaluation until judgement has been passed, which in our terms means until information has been related to a set of standards so as to enable the formulation of an evaluative statement, or a set of evaluative statements based on the simpler types of evaluative schemes.

Some of the most important sources of data on counselling were not, in fact, created under the formal auspices of the project. There were, for instance, two reports written by members of the Survey Research Department, and a short presentation of data by

the central Regional administration. However, if we restrict ourselves to the work conducted by the Tuition and Counselling Research Project, it seems that it either produced mainly descriptive reports or data which were not directly concerned with the effectiveness of the counselling service. We can consider the main reports produced by the project in turn.

One of the earliest of these was a report on the results of a questionnaire sent to a sample of part-time tutors and counsellors in 1971.<sup>6</sup> This summarised a considerable amount of information about the backgrounds and experience of tutors and counsellors, together with data on the tasks being carried out by them. It was largely descriptive, except for the conclusion that "the tuition and counselling system appears to be operating as planned",<sup>7</sup> which was obviously not intended at such an early stage to be a judgement of the 'effectiveness' of the services. It is perhaps worth noting in passing that although it was reported without comment that one Region seemed to have a high proportion of part-time staff who did not receive face-to-face briefing early in the year, the Regional Director concerned nevertheless wrote to challenge the figures.

As a result of the study centre visits project, which was the cause of such anguish as we have shown above, two reports were written. One of these consisted of a list of 'successful practices' which had been 'observed' by the participants as being carried out by counsellors. Its main purpose was as a source of information for briefing of Regional staff. The second report consisted of a collection of the reports prepared by the visitors to the centres. These included interviews



with students, tutors and counsellors, and descriptions of study centre facilities and activities. This was also envisaged as most useful for briefing and as a general source of insight into the workings of study centres. Neither report commented on the 'effectiveness' of the counselling service as a whole.<sup>8</sup>

One of the continuing objectives of the project was the monitoring of forms used by tutors and counsellors to record such things as their contacts with students, problems raised, grades, general comments on their work at the study centre, and so on. Some of these were intended mainly as useful in helping part-time staff keep track of their assigned students, whilst others were used to report problems to the Regional Offices and the central HQ, and to provide feedback to the policy making bodies. There were some difficulties in accessing some of these forms, partly because in these early days the channels for collating and feeding back the information had not been fully established. One form was, however, analysed by the project staff, with the assistance of Senior Counsellors, and a main report was issued early in 1972.<sup>9</sup> The form concerned was one that was used by counsellors to report on their general activities and problems at the end of the year. Information was derived on such topics as counsellors' methods of contacting students (telephone, home visits, etc.); types of activity during counselling sessions at the study centre (group discussion, individual counselling, etc.); ways used to encourage less forthcoming students to contribute to group discussions; problems envisaged in counselling in 1972; types of problems raised

by students; counsellors' views on their need for stronger links with other members of the University; and aspects of briefing and training which were missing or in need of modification. The report was a descriptive summary of the data, and did not draw conclusions about the 'effectiveness' of the counselling service.

In 1973, during the period in which the Review Committee on Tuition and Counselling was at work, the writer conducted the study of counsellor characteristics via the critical incidents approach, as mentioned above.<sup>10</sup> This involved obtaining from Senior Counsellors descriptions of behaviour of counsellors whom they regarded as either 'successful' or 'unsuccessful'. The outcome was a set of descriptive categories of behaviour which seemed to be associated with 'successful' and 'unsuccessful' counselling, together with a series of illustrative descriptions of specific instances as reported by the Senior Counsellors. It thus represented, in effect, an externalisation of the ordered criterion variables which entered into Senior Counsellors' evaluative schemes (Table 6). It was not, however, itself an evaluation of the counsellors, still less of the counselling service.

The other main sources of data on counsellors created by the project were those concerning the backgrounds of part-time staff, and the rates of drop-out and turnover within the population.<sup>11</sup> These were again descriptive.

**Table 6: Criterion Variables for Counsellors Identified by the Critical Incidents Method**

POSITIVE	NEGATIVE
<p>1 Shows enjoyment and enthusiasm for the job. Has a positive approach and regards the job as important.</p> <p>2 Shows initiative and independence in action. Is self-sufficient and accepts responsibility for helping students.</p> <p>3 Shows sympathy and empathy in relations with students. Is able to see situations from the student's point of view.</p> <p>4 Shows liking for and interest in students, and encourages contact.</p> <p>5 Has an ongoing manner and is relaxed, informal and confident.</p> <p>6 Keeps informed on the circumstances and progress of all students, and is able to respond positively to senior counsellor's requests for information.</p> <p>7 Identifies and pursues individual students' problems, giving practical constructive advice.</p> <p>8 Refers problems to senior counsellor when necessary, providing clear and detailed information. Considers and acts on advice given.</p> <p>9 Devises flexible programme of study centre activities. Prepares for sessions in advance.</p> <p>10 Creates and maintains group activities. Manages groups constructively and with tact, encouraging students' contributions. Creates relaxed, informal atmosphere.</p> <p>11 Shows knowledge of subject matter(s) and adapts approach to Open University context</p> <p>12 Knows administrative regulations and brings them to the attention of students when necessary.</p> <p>13 Co-operates with part-time colleagues. Seeks to co-operate and liaise with tutor(s), relating own 'tutorial' work to tutors where possible.</p> <p>14 Attends to management of study centre.</p>	<p>1 Is unenthusiastic about the job and does not derive enjoyment from it. Regards the job as a chore.</p> <p>2 Is passive and lethargic. Fails to accept responsibility for action.</p> <p>3 Lacks sympathy for students. Is unable to appreciate student's situation.</p> <p>4 Is remote, detached and uninterested in students and does not encourage contact.</p> <p>5 Is withdrawn, reserved, formal and lacks confidence.</p> <p>6 Does not keep informed on the progress and circumstances of students. Cannot give senior counsellor information when asked.</p> <p>7 Does not identify or pursue individual students' problems. Ignores problems or fails to give constructive advice.</p> <p>8 Passes all problems to senior counsellor or fails to consult senior counsellor at all. Fails to provide clear, detailed information. Does not respond to advice.</p> <p>9 Lacks programme of study centre activities. Treats session as disconnected parts. Works 'off-the-cuff' without preparation.</p> <p>10 Unable to initiate or maintain group activity. Dominates students and has patronizing approach. Does not encourage informal discussion.</p> <p>11 Shows ignorance of subject matter(s). Does not adapt approach to Open University context.</p> <p>12 Is unaware of administrative regulations and cannot advise students accurately.</p> <p>13 Ignores part-time colleagues and fails to liaise with tutor(s).</p> <p>14 Does not exercise responsibility for managing study centre.</p>

Although each variable is ordered 'internally', no order is implied among variables. The numbering 1-14 is therefore not intended to imply order.

Finally, a summary of data collected on counsellors and counselling was prepared in early 1973.<sup>12</sup> This drew together much of the data given in earlier reports, together with some collected by the Survey Research Department. No comment was made on the effectiveness of the counselling service.

Thus over the period considered, the project did not produce an evaluation of the counselling service, nor of counsellors. Given the current notions of evaluation held by evaluation practitioners, this is not perhaps surprising, since according to Stake most 'evaluators' have not regarded it as part of their job to produce evaluative statements. However, since as Stake points out, most clients expect the evaluator to make evaluations, which is not unreasonable, it is perhaps also not surprising that some parties were disappointed with the projects efforts in respect of counselling and the counselling service. There was, perhaps, an understandable mismatch of expectations.

Given that the project had as one of its main objectives the evaluation of the counselling service, and that it was not able to produce such an evaluation up to the time at which the counselling service became a major focus of debate in 1973, we can now consider some reasons for this state of affairs.

#### Difficulties in Evaluating Counselling

Firstly, it is important to remember that the researchers (including the writer) were not equipped from the start with a model of the evaluative process such as we have outlined. Nevertheless

there was an awareness of the need for some sort of standards for evaluation, and of the need to collect data on counsellors and counselling. We can thus consider the difficulties encountered in terms of the general framework which has been proposed.

#### Difficulties Concerning Standards

One of the first tasks required to be completed in the conduct of explicit evaluations is a definition of the subject and purpose of the evaluation en route to the creation of standards. One of the difficulties which occurred was that the subject of evaluation and its purpose were not made fully explicit. The initial brief referred to the evaluation of the effectiveness of the counselling service, which was in no way to involve a monitoring of individuals within the system. In terms of what can loosely be called the Faculty point of view, the purpose seems to have been to determine whether a service which they believed to be ineffective was in fact so, with a view to justifying their claim on these resources for further tuition. On the other hand, those more directly concerned with counselling seemed to regard the purpose of the evaluation as a means to establishing that the counselling service was satisfactory, and to identify areas in need of improvement.

The two orientations imply different types of evaluative schemes. In the first case, the actual subject of evaluation would have been the alternative of continuing the tuition and counselling arrangements as they were, as against redistributing some proportion of the counselling resources to other purposes such as tuition.

This would thus have involved at its crudest, the use of a criterion

variable such as 'student benefit'. It would have been necessary to estimate the value of this variable given by the existing arrangements together with a projected value as projected from the installation of a different system. This is, of course, loosely speaking what the Review Committee did when considering its various models of the tuition and counselling arrangements. Simply determining the value of a criterion variable for the existing counselling arrangements would only have provided half the story. There would still have been a need to project the value for alternative arrangements. Needless to say, such a projection would have been very difficult to make.

One possibility, raised in fact by a senior member of the Institute staff, was to conduct an experiment,<sup>13</sup> It seemed possible on prima facie grounds that a scheme could have been introduced whereby a number of Regions were matched and different tuition and counselling arrangements introduced into each. In theory at least this would have enabled the determination of whether different tuition and counselling arrangements made any difference. In practice, of course, such 'real-world' experiments are usually extremely difficult to carry out. Moreover, there were several reasons put forward against the suggestion. In particular, it violated the principle of equal provision for students. What would happen if under one scheme a high proportion of students dropped out or failed? Might not students in the experimental Regions feel they were being deprived of services their colleagues were getting? Would it be unfair to ask many counsellors who had given their all in the first few years

to stand down? What would happen if it was decided that they were wanted after all? And how could the administrative and information systems cope with a variety of different arrangements? In other words, would it all be worthwhile? The answer seems to have been that it would not.

The second orientation suggested a different kind of evaluative scheme. From this approach, the counselling service would be evaluated in terms of its objectives, with a view to identifying those which were or were not being achieved 'adequately' and hence initiating a search for remedies for the latter. The evaluative subject is then the counselling service rather than alternative courses of action.

Objectives are one of the traditional sources of standards for use in evaluation. Statements of objectives can be regarded as specifications of internally ordered criterion variables. In the case of counselling, there were two broad problems. One was that the objectives of counselling were not initially well established; also the officially stated objectives tended to be difficult to measure.

The specification of objectives for counsellors for 1971 said that "the main role of the Counsellor is to establish personal contact with the students assigned to him at a Study Centre .... and to guide, help and encourage these students in their studies."<sup>14</sup>

He was also to discuss course choices with students; encourage informal discussion groups and develop the "social educational side of the student's work";<sup>15</sup> help ensure the smooth running of the study centres; and possibly visit students in their homes and contact them by phone and by letter. There was no indication of how many counsellors would have to "guide, help and encourage" how many students for the service to be regarded as a success, and it is, of course, by no means obvious what would have counted as having guided, helped and encouraged.

Early in 1971, when the project was initially getting under way, Senior Counsellors were uncertain about objectives and standards. For instance, one group of Regional staff reported that it was "too early to decide on the degree of success of the counselling system."<sup>16</sup> Another group reported as follows:

"The counsellor's present role lacked definition which led to the possibility of dichotomous thinking among Senior Counsellors, some of whom saw the role almost exclusively in terms of the adult student learning in the adult education sense, and others who saw the need for a much deeper treatment of students' personal problems that might interfere with learning - this latter demanding greatly increased in-service specialist training."

"Meanwhile it was essential to specify objectives."<sup>17</sup>



Overall it was agreed that "no satisfactory definition [of success] could yet be formulated."<sup>18</sup>

By 1973 the objectives for the counselling service had become more wide ranging. The Assistant Director of Studies responsible for counselling wrote that:

"The major function of the counsellor remains to monitor and analyse factors which contribute to, and inhibit, the educational progress of individual students from his unique position of both thoroughly knowing the individual person, his interests, motives, attitudes and aspirations, and receiving all information on his assessment from course tutors and examinations."<sup>19</sup>

Increasingly important areas were giving advice to applicants, advising students on the selection of new courses, advising 'post-experience' students (those taking short courses and not registered as undergraduates), identifying potential withdrawal from courses by students, and 'caretaker' counselling (dealing with students during the November/December period at the end of the Open University's academic year, when study centres are closed). At this time it was possible to identify more than forty objectives and sub-objectives for counsellors, the most general of which included:

To guide help and encourage assigned students

To assist the student to act more effectively as a student

Not simply to be an extra source of subject matter information

To form and maintain friendly relations with students

To acquire a broad understanding of the course at Foundation level

To follow the progress of each individual assigned student

To check that arrangements at the study centre are working smoothly

To minimize students' feelings of isolation and depersonalization

To help the student understand the Open University system and the roles of Open University staff

It can be seen that even to determine whether counsellors had a 'broad understanding' of the course at Foundation level would have been a major undertaking, still less to measure the extent to which counsellors had minimized students' feelings of isolation and depersonalization.

We have noted that the formulation of standards can be a difficult task, and is one which is likely to involve a considerable amount of interaction between the researchers and the clients. One of the difficulties which faced the project was that it was not easy to determine who the clients were. The general client was the 'University', but a University is not an entity with which one can communicate. The 'parent' committee was the Tuition and Counselling Project Working Group, which became the Tutorial Board in the early days of the project's work. This body was responsible for formulating tuition and counselling policy and for controlling the tuition and counselling budget. As such it was perhaps the most relevant body for the setting of standards. Yet as has been said, the size and workload of this committee did not make it a forum conducive to the rather complex and difficult problems associated with the design of formalised enquiries. A continuing problem

for the project as a whole, as illustrated in the previous chapter, was that of obtaining coherent briefs for studies from committees such as the Tutorial Board. This was never fully resolved despite requests for clearer statements such as that put to the Board early in 1972. It thus seems unlikely that the Board would have been able to formulate a set of standards for counselling even if they had been specifically requested to do so.

#### Difficulties in Obtaining Information

Since clear standards and a clear definition of the purpose of the evaluation were not readily available, there was a problem of knowing what sort of information to collect. Even so, data on counselling was collected, but not always without difficulty.

There were two main sources of data which were tapped by the project.

The first of these was the report forms used by counsellors mentioned above. One of the aims of the project was to collate the information contained on some of these, although the central Regional administration took on more responsibility for this as time passed. The Institute's Survey Research Department took over the analysis of one of the forms in 1971.

The following forms relevant to counselling were considered for analysis by the project researchers.

Form C1:- this form was intended to be used by counsellors to record contacts and problems raised by students. They were supposed to be completed over the year for each student assigned to a counsellor. There seemed to be a problem in obtaining these from counsellors efficiently, and as the information contained on the form was summarised on another (C3, see below), they were not obtained for analysis.

○ Form C2:- this was to be used by counsellors to refer urgent problems to Senior Counsellors. It was thought that an analysis of the forms might indicate the range of serious problems counsellors' were encountering. One set of these was analysed early in 1972. They were felt to be somewhat misleading since Senior Counsellors on the project Planning Group said that most counsellors used the phone to report really urgent or serious matters. On the advice of the Planning Group, no larger scale analysis was conducted.

○ Form C3:- this summarised the contacts and problems raised by students with counsellors, and was completed by them for each assigned student. The 1971 forms were analysed by the Survey Research Department. For 1972, the central Regional administration took over the analysis and produced a short report.<sup>20</sup>

Form C4:- this was used by counsellors to provide a general report on activities and problems at the study centre. They were analysed by project staff for 1971 with the assistance of Senior Counsellors. They were modified for 1972, and were analysed by Regional staff and the central Regional administration.

Table 7: Report and Record Forms Used in the Counselling System \*

Title	Content and Purpose
C1	Completed by counsellor. One form for every assigned student. Record each contact and the nature of the problems raised. Help counsellor keep track of students; aid to continuity if student changes counsellor.
C2	Counsellor's report students' problems to Senior Counsellor, who replies on the same form. General purpose counsellor - Senior Counsellor communication medium.
C3	Completed by the counsellor at the end of each year. One form for every assigned student. Summary data on the number of contacts, channels used (e.g. telephone, home visit, etc.), types of problems raised.
C4	Completed by the counsellor at the end of each year. Provide feedback on counselling activities and study centre operations. Wide range of questions on counselling, including patterns of activity at counselling sessions, counsellor's use of time, most important matters raised by students, etc.
C5	Report card sent by a counsellor to his Regional Office giving the number of students attending a session. Could be used to monitor attendance.

\*A number of other forms were in use, but these were mainly for 'special' reports, e.g. on disadvantaged students.

Form C5:- this was used only in some Regions. It was used by counsellors to report the numbers of students attending sessions, and contacts by other means. Analysis of this was largely in the hands of the Regions.

Generally the analysis of these forms passed into the hands of those outside the research group after 1971. There were generally acknowledged to be problems with the reporting system used by part-time staff, and major revision was planned for the introductions of the revised tuition and counselling arrangements in 1976. Some of the early problems were that some of the forms were of poor design, making the interpretation of responses difficult; the machinery for gathering together the forms was not well established or was non-existent; and there was uncertainty about who had the responsibility for conducting the analysis. These factors, together with the limited resources of the project, prevented the researchers building up a body of feedback on the basis of these forms.

There was in any case considerable difficulty in interpreting the data which was derived from these forms. The report on the analysis of the C3 forms for 1971 said:

"When it comes to interpreting these tables, however, we are met by certain difficulties. These mainly concern the reliability and validity of the information given by counsellors .... Counsellors vary as to what is considered a mentionworthy problem and how to classify such a problem. They vary in their working definition

of a contact and the degree to which they expand their responses with written comments. If a certain problem is not recorded we do not know whether no problem existed or whether the counsellor was merely not aware of it. Many questions remain unanswered and although we can assume that if the information was available it would not affect the results, this may not be the case."<sup>21</sup>

Similarly, a report on an analysis which covered the C5 form data for January to March 1973<sup>22</sup> was prefaced by seventeen paragraphs on the 'limitation of the data'. These included very small samples; problems of knowing what a 'contact' might mean; doubts about whether respondents might have forgotten about some of the 'contacts'; doubts about whether some of the respondents might have "swelled their statistics from time to time"; inability to determine whether respondents were 'contacting' all their students or not; difficulty in determining whether 'contacts' were with Foundation level or post-Foundation level students or even students assigned to another counsellor. The Director of Studies, Local Centres and Tutorial Services concluded that:

"I think it is clear that forms which were chiefly designed as part of the regional process of contact and supervision between part-time staff and regional academic staff reveal very little of wider significance - other than that contact between counsellors, tutors and their students proceeds with some apparent (though not significantly established) variations across the UK."<sup>23</sup>

A second main source of data on counselling was that collected by the Survey Research Department as part of their programme of student-based enquiries. Although this data was thus neither collected nor analysed by the project researchers, it constitutes an important source of data on counselling and the counselling service. It can thus be considered in the light of the general difficulty of interpreting data on the counselling service.

The data was obtained from a number of questionnaire surveys of students in 1971 and 1972.<sup>24</sup> The 1971 material was mainly concerned with two groups of students, those who did not continue to study beyond the first three months of the year, and those who either stopped studying before completing the course or who completed the course but did not continue into the following year. The surveys of these two groups of students received 66% and 62% response rates respectively.

Taking the first group, known as 'early drop-outs', the data (Table 8) showed that about a third attended their study centres 'less than once a month' or 'never' on average. Sixty-five per cent said they attended 'about once a month' or more frequently. When those who attended less frequently were asked for their reasons (Table 9), the largest response (25%) was in terms of 'work commitments'. None of them gave 'inefficient counselling' as a reason. When asked about their satisfaction with the counselling arrangements (Table 10), 15% thought they were less than 'fairly satisfactory' in comparison with 30% who felt that way about the class tuition. When they were asked whether they had asked for 'help with the course materials'



(Table 11), only a quarter had, and the counsellor seemed to have been more helpful than the class tutor, even though at that time he did not have a 'defined tutorial role'. On the other hand, when it came to the question of deciding to withdraw (Table 12), 25% 'discussed' it with their counsellor, 32% with their family and 37% claimed not to have 'discussed' it. To put it another way, 75% seemed not to have 'discussed' it with their counsellor.

For the group who continued beyond the first three months of the year but who then either dropped out or did not proceed to 1972, about a third attended the study centre 'less than once a month' or 'never' on average (Table 8). Again none of them mentioned 'inefficient counselling' as a reason for this, although 6% did mention 'inefficient tutoring' (Table 9). However, there was clearly some difficulty in interpreting this, since the report said that "this does not necessarily mean that counselling was 'entirely good' but that tuition was bad enough to cause non-attendance for 6% of these students."<sup>25</sup> Presumably 'inefficient counselling' could not be regarded as unconnected with the less frequent attendance of these students at study centres, even though none of them actually gave it as a reason. Presumably it would have been equally possible to argue that even though 6% had mentioned 'inefficient tutoring', this did not in fact affect their attendance. Fifty-seven per cent of this group, however, were 'very satisfied' or 'fairly satisfied' with the counselling service (Table 13), although between thirty and forty per cent seemed either not to find the service satisfactory or not to use it.

Table 8: Frequency of Attendance at Study Centres Claimed by Early Drop-outs and Non-continuers, 1971 (percentages)

Frequency of Attendance	Early Drop-outs	Non-continuers
	%	%
Weekly or more often	19	18
About once a fortnight	29	27
About once every 3 weeks	8	10
About once a month	9	14
Less than once a month	13	22
Never	20	9
No answer	3	2
Total respondents	759*	1014

Source: N.E. McIntosh & S. Blacklock, Students and Counselling, May 1973, p. 8.

\*Excludes those not assigned to a study centre.

Table 9: Reasons for Low or Non-attendance at Study Centres by Early Drop-outs and Non-continuers, 1971 (percentages)

Reasons	Early Drop-outs	Non-continuers
	%	%
Not useful to attend	10	24
Work commitments	25	16
Travelling/expense difficulties	10	14
Domestic/personal inconvenience	14	11
Prefer to study on own	-	11
Not enough time	8	7
Inefficient tutoring	0	6
Too busy with course	-	4
Illness	6	3
Lack of information on where and how to attend	5	3
Inefficient counselling	0	0
Other reasons	11	5
No answer	18	10
Base	272	309

Source: as Table 8, p. 10.

- less than 0.5%.

Table 10: Satisfaction with the Tuition and Counselling Arrangements Claimed by Early Drop-outs, 1971 (percentages)

Degree of Satisfaction	Counselling	Class Tuition	Correspondence Tuition
	%	%	%
Very satisfactory	25	9	21
Fairly satisfactory	31	27	26
Not very or not at all satisfactory	15	30	11
Not used	12	24	28
No answer	17	11	14
Base	883	833	833

Source: as Table 8, p. 11.

Table 11: Sources Asked for Help with the Course Materials by Early Drop-outs and Whether Found Helpful, 1971 (percentages)

Sources Asked	Asked for help	Found helpful
	%	%
Class Tutor	52	27
Correspondence Tutor	5	3
Counsellor	49	31
Staff Tutor	4	3
Senior Counsellor	2	2
Regional Office	9	2
OU Headquarters	7	0
Other students	32	22
Base = those asking for help	207	207

Source: as Table 8, p. 16.

**Table 12:** Persons with Whom Early Drop-outs and Non-continuers Discussed Withdrawal, 1971 (percentages)

	Early Drop-outs	Non-continuers
	%	%
Class Tutor	3	4
Correspondence Tutor	0	2
Counsellor	25	31
Staff Tutor	0	0
Senior Counsellor	1	2
Regional Office	3	5
OU Headquarters	2	4
Other students	11	14
Colleagues at work	15	16
Family	32	42
Friends	14	15
Did not discuss it	37	37
Base	not available	not available

Source: as Table 8, p. 18.

**Table 13:** Satisfaction with the Counselling Service Claimed by Non-continuers, 1971 (percentages)

Degree of satisfaction	Non-continuers
	%
Very satisfactory	28
Fairly satisfactory	29
Not very satisfactory	14
Not at all satisfactory	7
Not used	16
No answer	7
Base	1014

Source: as Table 8, p. 13.

Table 14: Satisfaction with the Tuition and Counselling Arrangements Claimed by Students Completing Courses in 1972 (percentages)

Degree of satisfaction	Counselling %	Tuition %
Very satisfactory	34	22
Fairly satisfactory	37	41
Not very satisfactory	11	21
Not at all satisfactory	7	9
Don't know - never use	9	5
No answer	2	2
Base	1135	1131

Source: as Table 8, p. 19.

Table 15: Main Reasons for Dissatisfaction with Counselling Given by Students Completing Courses in 1972 (percentages)

Reasons	%
No help, not able to give help, give information	17
Needs more specialised knowledge about courses in order to be able to advise	12
No/unsatisfactory help for 2nd level students	11
Counsellors unsure of their role, no real sense of direction	11
Difficulty in contacting counsellor	8
Teaching by counsellor not up to required standard	8
Difficult to see counsellor alone	6
Difficulty in attending sessions	7
Don't need/want counsellors	5
Base = all not satisfied	205

Source: as Table 8, p. 18.

Sixteen per cent of this group said they 'did not use' the counselling service, but in response to a question on whether they ever contacted their counsellor, 28% indicated they had not. Those who did contact the counsellor seemed to have done so for reasons which were not primarily course-based. Thus "on the face of it the counsellor seemed to have been fulfilling the function laid down for him."<sup>26</sup> But among this group, 31% had discussed withdrawal with the counsellor, 42% with their family, and 37% had not discussed it. (Table 12).

Some information was also available from 1972 students. Overall, students were rather more happy with the counselling than with the tutoring arrangements (Table 14). Again there were difficulties in interpretation. The report said that "it is difficult to avoid the thought that this may be a negative satisfaction, arising from lower demands, and therefore lower expectations. Since many students have used counsellors so little, they may have little cause to find them unsatisfactory."<sup>27</sup> Yet only 18% said they found the counselling arrangements less than 'fairly satisfactory', and only 9% said they 'didn't know' or 'never used it'. Moreover, only 5% of those who had said they found the arrangements less than 'fairly satisfactory' said that they 'didn't need or want' counsellors (Table 15). Thus students seemed by and large to be satisfied, but it seemed to be hard to believe that they 'really' were, or that if they were satisfied, they were satisfied for the 'right' reasons'.

It is not intended to suggest that the interpretation made by the reporters was necessarily incorrect, nor that there was necessarily

nothing 'wrong' with the counselling service. It is simply intended to illustrate that such data is often difficult to interpret, and is often open to questions concerning validity. As Phillips says:

"What I have been arguing is that much of the evidence used by sociologists is of questionable validity; that furthermore, the relationships between our chief independent and dependent variables are generally quite weak; and, finally, that even if we had valid measures and strong relationships the problems of interpretation are formidable."<sup>28</sup>

Thus, as the reporters indicated, many questions needed to be answered before the data could be adequately interpreted.

The problems in evaluating the counselling services were thus many sided. There were difficulties in arriving at a clear definition of the problem at hand. There were difficulties in arriving at standards and doubts about the objectives of the counselling service. Data were not easy to obtain, sometimes because of anxieties about how they might be interpreted or of the disruptions which might be caused in obtaining them. And the data which were obtained were sometimes difficult to interpret and open to argument concerning validity.

How could these problems have been solved? The answers must be somewhat speculative. The research-type problems are not simply ones which arose in this particular context. They are faced by the social sciences as a whole. There does not seem to be any fundamental solution at hand. Perhaps all that can be hoped for is an acknowledgment by the parties concerned that it is probably most unlikely that

they will ever be certain of 'what is really going on'. Instead it might be useful to try to reach an agreement on what is going to be accepted as evidence before the evidence is collected.

On the question of standards, it is obviously helpful to have some idea of what a set of standards might look like, and the proposed model attempts to give some indication of this. It certainly is important to know whether the world is as it ought to be, but if we want others to find out for us we must be able to say how it ought to be. If this is not done, at best the evaluator will tell us whether it is as he thinks it ought to be, or as what he thinks we think it ought to be.

Evaluative disputes are often bathed in emotion. We tend to regard those who seem not to share our standards as not only mistaken but perhaps wilfully obstructive or even 'bad' people. Such was the character of the religious wars of medieval days. One's opponents were not simply mistaken; they were evil. As such there was little point in listening to them, still less in believing that they might hold perfectly justifiable views. There was thus no point in trying to understand what those views were. It is understandable that we should have strong feelings about the things we hold dear, but it is equally true that they can lead us into an unhealthy rigidity and obscurantism in our thinking. At the same time, it can lead to a good deal of interpersonal acrimony, which, in the writer's view, can be decidedly unhelpful in grappling with the real problems which are of concern to the parties.

It is not possible to say whether such problems can be overcome. But it would seem to be a very great advance if the



Figure 13: Some Hypothetical Standards for the Counselling Service

- If 90% or more of students say they are very or fairly satisfied with the counselling service, then the service is very satisfactory
- If 50% - 89% of students say they are very or fairly satisfied with the counselling service, then the service is fairly satisfactory
- If less than 50% of students say they are very or fairly satisfied with the counselling service, then the service is unsatisfactory
- If more than 50% of counsellors have more than 50% attendance at more than 50% of their counselling sessions, the service is successful
- If 50% or fewer counsellors have more than 50% attendance at more than 50% of their counselling sessions, the service is unsuccessful
- If 90% or more Senior Counsellors say that the counselling service is successful then it is successful
- If 30% - 89% of Senior Counsellors say that the counselling service is successful, then it is fairly successful
- If less than 30% of Senior Counsellors say that the counselling service is successful, then it is unsuccessful
- If more than 50% of counsellors say they dislike counselling, then the counselling service is very unsatisfactory
- If 10% - 15% of counsellors say they dislike counselling, then the counselling service is fairly unsatisfactory
- If up to 10% of counsellors say they dislike counselling, then the counselling service is satisfactory
- If no counsellors say they dislike counselling, then the counselling service is very satisfactory

assumptions, beliefs, theories and standards which are used by the parties could be made explicit. At least then the basis for argument would be clear.

#### Additional Difficulties of the Project

There was a recognition by the researchers that the topic of counselling was a 'sensitive' one. As has been noted, there were strong feelings amongst interested parties about the value of the counselling service. In particular, those who were most directly concerned with operating the counselling service were perhaps most sensitive about it, and there were probably a number of reasons.

Firstly, there was some suspicion about the motives which underlay the conduct of the project. In the view of some of the 'supporters' of counselling, the research, or the idea of 'evaluating' counselling, was regarded as perhaps part of some Machiavellian scheme designed to attack what they regarded as an important and beneficial service to students. They were thus not necessarily confident that whatever data were collected would be interpreted in an 'impartial' manner. Since the standards to be applied were not clear, this was understandable. Requests for data were thus sometimes able to be dismissed as attempts to pander to the desires of the 'opponents' to 'dig a hole' under the counselling service.

This atmosphere was probably not helped by the fact that the project was located in the Institute of Educational Technology.

As an important body in the University, the Institute was naturally

expected to make known its views on various aspects of University policy, including the policy for tuition and counselling services. As we have seen, the Institute's philosophy was broadly opposed to the provision of 'personal' tuition and counselling. The Director of the Institute was in a particularly difficult position as both head of the Institute and of the research. Thus it was hard for the project to maintain an image of impartiality when the Institute was known to be generally 'against' the local services, and in particular when it advocated halving the counselling budget. The impact of these views on the actual attitudes of the researchers is perhaps less important than the interpretation which could be put on them by those whose cooperation was being sought. It could easily be viewed as confirmation of their worst fears. Thus for instance, the writer was asked at a meeting of senior Regional staff from three Regions in 1973, whether the research into counsellor characteristics was intended to help counselling, or whether it was really intended to help "the friends of IET". Similarly, one Region refused to allow one of the researchers to interview one of their counsellors for what one suspects were similar reasons.

Secondly, there was sensitivity about research into counselling because of doubts about the capacity of research to obtain more than a superficial impression of what counselling was supposed to be about. There was thus some feeling that counselling was an intensely personal and in some ways private activity which could not be captured by questionnaires and statistics. There was some feeling that the achievements of counsellors could not be easily quantified, and that

the kind of 'global' research which the project would have conducted would only produce superficial data which would nevertheless be used to attack the service. An analogy might be drawn with the current 'attacks' on local authority social services:

"For any Town Hall treasurer hunting savings, the personal social services are a winsome target. Social workers have most difficulty of all in quantifying their work. Unlike housing, with its statistics on families accommodated, or education with its numbers of children taught, the social workers have no way of calculating the number of broken homes or battered babies they may have saved."<sup>29</sup>

Similarly the counselling service had no obvious way of determining how many students had been prevented from dropping out by counsellors, or the extent to which students had been helped in their 'social educational' or personal development. Even a measure of contact between counsellors and students, which at least seemed to be an indication of how many students might have benefited from counselling, could be held to be an inadequate measure of the 'success' of the counselling service, since it could be argued that students benefited from knowing that there was a counsellor to go to even if in fact they did not. Thus the argument was like that which could be applied to high-wire artistes; the fact that many do not fall into the safety net is no grounds for saying that it is unnecessary, because it is the presence of the net which stops them falling. In the absence of experimentation, it is difficult to establish the truth of this,

and as has been noted, some thought that such an experiment in respect of counselling would have been no less dangerous than one with high-wire walkers.

A third source of sensitivity which can be suggested was that those responsible for counsellors were worried about the effect of inquiries into counsellors' work on counsellors themselves. They were already thought to be over burdened and in some ways going beyond 'the call of duty' for the University. Moreover such enquiries might be thought to be contravening the traditional taboos concerning the 'autonomy of the teacher' or representing some sort of challenge to the professional status of University colleagues. It was also not clear what demands in terms of participating in research studies could legitimately be made on part-time staff. An illustration of the awareness of the need to take professional sensitivity into account is given by the fact that in the preparation of training materials for part-time staff, it was decided to avoid the use of the word 'training' since this might give the impression that part-timers were not fully competent.

As was noted in the case of the 1971 visits to study centres by senior Regional staff, some of the participants displayed an unexpected degree of sensitivity. There were no doubt various reasons for this, but the writer can offer one possible explanation. Since the Open University was a completely novel institution, it involved the creation of new roles and new organisational forms which were without counterparts in the wider educational system. Originally there were to be no Faculties, but 'lines of study'. There were to

be Directors of Studies rather than Deans. There was to be a Finance Office rather than a Bursary. And there were to be the roles of Senior Counsellor and Staff Tutor, among others. These roles, by virtue of their novelty, did not have any immediately recognisable prestige value, yet they were filled by persons coming from positions in the wider educational system which had established levels of prestige and authority. The incumbents of the new Open University roles could thus be seen as trying to establish the prestige and scope of their roles in this new situation. This emerged in some cases in terms of the flexing of muscles when others seemed to be according the role less prestige or authority than the incumbents thought was legitimate. Hence, for instance, the resentment of some Staff Tutors who thought they were being treated as lowly IET research assistants during the study centre visits project. This concern to establish the role was perhaps most acutely felt by Senior Counsellors, who were often the 'gatekeepers' to counsellors. Their position was particularly vague in the early days, the job seeming to have only a limited academic element and without clear lines of development inside or outside the University. They were thus likely to be particularly concerned to establish the scope and prestige of their role.

One of the ways of changing the prestige of a role is to change its name, for example 'rat-catcher' to 'rodent operative', or 'labour exchange' to 'employment office'. It is perhaps supportive of this thesis about the concern for establishing prestige and authority in an 'anomic' situation, that within a few years of the University's

start, 'lines of study' had become 'Faculties', 'Directors of Studies' had mostly become 'Deans', and the 'Finance Office' had become the 'Bursary'.

### Summary of Problems

We have considered, then, some of the difficulties associated with the evaluation of the Open University's counselling services.

We can briefly summarise these as:

1. Although there appeared to be an initial consensus between the researchers and the 'clients' over the nature of the research, the subject and purpose of the research were not definitely established and hence nor was the evaluative scheme. This was partly due to disagreement among the 'clients' on these matters.
2. The counselling service, by its very nature, was difficult to investigate rigorously, due to the rather vague nature of its objectives, and the difficulty in establishing 'facts' about possibly crucial matters such as the extent to which a student had been helped by a counsellor.
3. The 'political' context of the project meant that participants were not always willing to cooperate in securing data. This was perhaps worsened by the fact that there was plenty of 'informal' evaluation taking place, in terms of views and opinions on what was happening and how 'satisfactory' it was. In such circumstances, the introduction of 'formal' evaluation and research can be interpreted as implying that those who have given their honest views, based on their own experience, are not to be believed, or, as we have seen in the case of the visits to study centres, are incapable of making judgements of any worth.

4. The official 'feedback system' of forms used by counsellors did not always enable the right sort of information to be obtained in respect of the 'effectiveness' of the counselling service. The information was usually open to questions of validity and problems of interpretation.
5. The limited resources and 'semi-official' status of the research group, and uncertainty about who had the main responsibility for securing data, reduced the projects capacity to produce a comprehensive piece of work on the counselling service. The wide remit of the project, with its consequent involvement in topics such as summer school evaluation and rates of turnover among part-time staff, inhibited a concentrated approach on the topic of counselling.

Since the project seemed to be expected to cope with many different problems arising from different parts of the University, and since it tried to cope with a considerable number of these, there was less time and resources to devote to single topics. We have seen, however, that there was an increasing concentration of effort as the years progressed.

Some of these problems undoubtedly arose because of the novelty of the University and the infant stage of its development. The system of forms for use by part-time staff, for instance, is being revised for the introduction of the new tuition and counselling arrangements in 1976. Also in the early days, there had been no opportunity to build up personal relationships between researchers and participants, and the dispersed nature of the University



(with researchers and Faculties at HQ, and regional staff dispersed throughout the country) meant that there was relatively little opportunity for informal interchanges. In time it was possible to get to know the people who lay behind the memoranda and to build up a measure of goodwill towards the research work. This was most important.

When the new tuition and counselling arrangements are put into effect in 1976, it is intended that they should be 'evaluated'. There is no way of knowing what form this 'evaluation' will take, or what its purpose is intended to be. Whatever the case, there will no doubt be problems. Certainly if this is to proceed in a coherent and formalised way, there will need to be much preparation. Perhaps we have been able to offer some insight into what needs to be done if an evaluation is required in the sense which we have established, and of the problems which are likely to be encountered along the way.

#### Some Comments on Evaluation and Decision-Making

One of the main roles which is normally seen for formal evaluation is to improve decision-making. The logic is impeccable: some state of affairs is desired: - first see if the desired state of affairs already exists (evaluate the current state of affairs) since it would be foolish to try to produce something which is already there:- if the desired state of affairs is not current, identify a set of courses of action likely to change it and select the best (evaluate alternative courses of action). For reasons we have already given, even the empirical aspects of this process

in respect of the social world are difficult to execute. A graver point, however, is that even the most formal and sophisticated methods for obtaining and handling information are of very little pragmatic value if decision-makers do not in fact allow them to influence their decisions.

One of the main assumptions which underlies the advocacy of formal methods of decision-making (for instance by using operational research procedures to generate possible courses of action) seems to be that once the decision maker has been persuaded to define the decision situation, specify standards, and has been provided with valid information, he will then make a 'rational' decision in terms of the explicit model which has been created. It is thus assumed that the decision-maker either is wedded to the same 'rational-empirical' philosophy as the researcher, or that at least he ought to be wedded to such a philosophy. There is, however, no guarantee of this, the implication being that in this case there seems to be very little point in wrestling with all the formal procedures.

If common sense always prevails in the end, why bother to go beyond common sense in the first place? Certainly it seems that it is indeed most important to go beyond common sense, to secure the best quality information and to make decisions in the most rational manner possible. Yet it is as well to be aware that the evidence suggests that this is the exception rather than the rule. The 'rational-empirical' philosophy envisages a sequence of formulation of decision situation, input of information, decision dependent on information. A more realistic sequence might be decision, input of information, construction of decision situation via selective

use of information, decision independent of information. In the former case, the decision-maker uses the information as a basis for his decision; in the latter, he uses some of the information to justify a decision already made on common-sense grounds.

To take some empirical examples, the question of whether it is better to have smaller class sizes in schools from the point of view of students' achievement, has long been a subject of debate. The relation of class size to achievement has also been the focus of a great deal of research. Summarising over 200 studies of the subject, Rossi wrote:

"By and large class size has no effect on the learning of students with the possible exception of classes in the language arts. But the net results of more than two hundred researchers on educational ideology and policy have been virtually nil. Every proposal for the betterment of education calls for reductions in the size of classes, despite the fact that there is no evidence that class size affects anything except possibly the job satisfaction of the teachers."<sup>30</sup>

Later studies by Husen, Little and Russell, Morris, and Davie<sup>31</sup> even suggested that students achieved more in larger classes. Nevertheless, reduction in class size remains a priority for at least the teaching profession. Now it may well be that class size is related to, say, teacher job satisfaction. Why not then use this as a criterion variable rather than achievement which is shown

not to be related to class size? Is this perhaps the real criterion variable which is being used? If so, it is not surprising that the studies have little to say about its relation to class size, since it is not acknowledged as a criterion variable. If the teachers are not prepared to admit, perhaps even to themselves, that this is their criterion, the only way to maintain the evaluation of class size is by ignoring the evidence, or, more probably, attacking its validity. Moreover, even if they were prepared to admit privately that this was their criterion, they are now in no better position to make a rational judgement since the studies have not necessarily been measuring this variable. Perhaps this is one reason why decision-makers seem to make decisions in spite of the formal evaluation. They only include in the standards, where these are explicit, variables which they regard as 'respectable'. Since the real criterion variables are not included in the standards, it is not surprising that the evaluations based on them are ignored.

There is, for instance, a well known story (perhaps true, perhaps a myth) in market research circles about two motor manufacturers who were each planning to introduce a new model.<sup>32</sup> The first company sent out its market researchers who asked people what sort of a car they wanted. They said they would like something economical, with plenty of room for the family, easy to drive and park, modest and not ostentatious. The second company told its market researchers to ask people what they thought their neighbour would want from a car. They were told it should look fast and impressive, have lots of chrome and fancy gadgets, and make no concessions to safety for the

sake of looks. The first company, it is reputed, made the car everyone said they wanted and lost several hundred million dollars. The second made a handsome profit.

The point about this is not that decision-makers necessarily cynically misrepresent their standards. The respondents to the first market research team no doubt believed everything they said, at least while they were saying it. It does, however, suggest that getting at significant criterion variables is not necessarily a straightforward matter.

Pursuing the question of class size and achievement, according to Shipman,<sup>33</sup> the Plowden Committee was faced with its own survey on the relation between class size and achievement, which concluded that larger classes seemed to facilitate greater achievement. The Committee's conclusion, however, was that reduction in class size should remain a priority. It justified this by saying the research results were outweighed by professional advice, public opinion and the example of other countries. "The writers of the Report", says Shipman, "seemed to have used evidence only where it supported their views and explained it away where it opposed them ..... the evidence which opposed or failed to support the recommendations ... was ignored or described as inexplicable and opposed to common knowledge."<sup>34</sup>

Carter<sup>35</sup> has also written about the reaction of decision makers to specifically negative findings in evaluation studies.

These seem particularly important since, according to Weiss

among others, "competent evaluations have come out with negative results in field after field."<sup>36</sup> Carter notes the problem of adequately defining evaluative research, but refers to Suchman's definition, noted earlier. By negative findings, he means "research results and conclusions that are opposed to the expectations of the client for whom the research was conducted."<sup>37</sup> Among the cases cited are the following:

CASE 1

A manager introduces an 'employee motivation program' into a plant. He reports to his immediate superior and to other plant representatives at an early stage that the program is a success. Six months later, a formal evaluation of the program is conducted. The evaluator obtains data on employee motivation from the manager for the period preceding the introduction of the program. He gathers further information on employee motivation after the adoption of the program.

A comparison shows that the program had not produced the expected results, and contradicted the manager's judgement. The manager refuses to provide clarificatory data, and wants no one else to see the evaluator's report. The program continues unaltered by the findings, and the relationship between the manager and the evaluator becomes cool and impersonal.

CASE 2

A top executive in a company makes a number of changes in operating procedures and management-employee relationships in a plant to try to improve production. Striking improvements follow. The executive then presents a similar scheme to a number of other plants.

Within two years the scheme has been dropped in the first plant, and the attempts to introduce in into other plants have failed. The scheme is nevertheless offered commercially to other organisations, and the executive ignores the fact that there are no data to indicate its long-term effects.

CASE 3

A highly respected survey research organisation in the United States conducts an investigation into the effect of fellowships and scholarships on students' choices of discipline for graduate study. The providers of these awards believe that they influence students to pursue the disciplines which are allocated the grants. The results show that they have little effect on students' choices of discipline. The relationship between researchers and clients becomes cool if not distant. The client continues to demand more money from the government on the grounds that the awards are, among other things, helping to attract the better students to the disciplines which offer them.

CASE 4

A sociologist is called into study communication and coordination processes in an administrative division of a large university. He is asked to submit recommendations for the improvement of these processes.

As a result of his study, a number of problem areas are identified. Action is taken by the client on those matters which were originally thought to be troublesome; other problems revealed by the study tend to be ignored. The researcher concludes that "for whatever reason, action research in this type of setting may prove simply a way of legitimating the pre-conception of organisational leaders."<sup>38</sup>

Carter notes some consequences of this state of affairs. One is that there seems to be a danger that once respondents in evaluation studies begin to realise that the results are selectively interpreted, they will cease to cooperate. Moreover those responsible for providing information to decision-makers from within an organisation may only supply their superiors with what they think they want so as not to become identified as 'trouble-makers'. This reflects, in the writer's view, the curious phenomena whereby researchers who produce results which clients do not like tend to become blamed as if they had themselves actually created the situation which they describe. The deteriorating researcher-client relationships which Carter reports as following



the communication of negative findings seems to provide an illustration of this phenomena.

If we examine the cases reported by Carter together with that of the Plowden Committee's handling of the question of the relation between class sizes and achievement, we find that they share a number of elements. In each case, the client has had an initial brief about some state of affairs. In each case, research evidence has contradicted this belief. In each case, the client has rejected the evidence rather than modify his belief. This is particularly evident in the case of the Plowden Committee, which initially believed that larger classes were associated with lower achievement. The existing research evidence seemed either to give no support to this belief or tended to indicate that its opposite was true. The Committee commissioned its own study, which failed to support its belief. Rather than change its belief, the Committee rejected the research.

In Carter's second example, the manager is unable to accept the evidence concerning the effects of his productivity scheme. In the third example, the providers of the scholarships behave as if the evidence produced by the survey research on the effect of the awards on students' choices of disciplines did not exist. In the fourth example, the University administration behaves as if only those problem areas which it initially believed to exist had been revealed by the research.

In each of these cases the clients behave in an incorrigible manner. It seems, in fact, as if their informal evaluations are based on incorrigible propositions.

The term 'incorrigible proposition' was coined by Gasking to refer to propositions "which you would never admit to be false whatever happens."<sup>39</sup> A corrigible proposition, on the other hand, "is one that you would withdraw and admit to be false if certain things happened in the world."<sup>40</sup> For example, the proposition 'I exist' became an incorrigible proposition for Descartes, since it was the one proposition he found himself unable to doubt; he could not conceive of anything happening that would make him believe this was not true. However, the proposition that 'I was born in 1948' is corrigible, since I would acknowledge it as untrue if my parents denied it, my relations denied it and my birth certificate showed a different date.

Incorrigible propositions tend to be more easily identified when we are dealing with cultures other than our own. The anthropologist Evans-Pritchard's experiences of the Azande Indians, with their belief in oracles, provides a further example.<sup>41</sup>

The Azande consult the oracle to obtain answers to important questions. They may, for instance, ask whether a sick person will live. To obtain an answer, a chicken is given a dose of what Western science terms 'poison', and the subsequent death or survival of the chicken corresponds to an affirmative or negative answer to the question on the part of the oracle. Since the events which the oracle predicts occur independently of the predictions, Evans-Pritchard is

led to ask how the Azande's belief in the oracle is maintained. He thus presents to the Azande the argument that since the oracle's predictions are wrong as often as they are right, this suggests that no oracle exists. The Azande, however, either refuse point-blank to countenance the possibility that oracles do not exist, or they present explanations which enable them to reconcile the failure of the oracle's predictions with its existence. Evans-Pritchard calls these explanations 'secondary elaborations of belief'.<sup>42</sup> Thus they might say that witches and sorcerers sometimes intervened to reverse the oracle's predictions. The possibility that oracles do not exist is simply not countenanced. For the Azande, belief in the oracle is incorrigible.

With these notions in mind we can see the behaviour of the Plowden Committee in rejecting the research evidence, as like the Azande rejecting evidence about the failings of the oracle's predictions. For the Azande the necessary secondary elaborations of belief are in terms of propositions about witches and sorcerers; for the Plowden Committee, their elaborations imply that something must have 'gone wrong' with the research. They behave as if their belief in the relation between class-size and achievement is incorrigible, insofar as they ignore or explain away evidence to factual propositions.

If we limit ourselves for the moment to factual propositions (i.e. empirical rather than superempirical propositions) the most commonly accepted test of the truth of these in our society is by reference to experience. A proposition is made about some aspect of the world, the world is then scrutinised, and a proposition is

formulated on the basis of this scrutiny. If the first proposition (belief) matches the second (knowledge) the propositions are true. If they do not match, either the first proposition is false (the belief was incorrect) or the second proposition is false (the 'knowledge' is invalid). For example, I may say 'There are three chairs in the room'. If I then count the chairs in the room I either conclude that 'There are three chairs in the room', in which case I decide my initial belief was correct; or I conclude that 'There are not three chairs in the room'. In this case I either decide my initial belief was incorrect, so really there are only two chairs; or I decide my second proposition is false. How can this be? I have just counted the chairs and there are only two. Perhaps I miscounted. Perhaps I said to myself 'one, three, two' as I counted the chairs. Perhaps my eyes are 'going funny'. Perhaps one chair was hidden behind another so I missed it. In other words, perhaps the conditions for valid proposition formulation were not met.

If my belief in the existence of three chairs in the room is incorrigible, then my only recourse when presented with evidence to the contrary is to establish the invalidity of the evidence. Of course it is possible simply to refuse to believe the evidence, to dismiss it point-blank. In general, however, sheer obscurantism tends to be regarded as unacceptable, and in extreme cases as pathological. In the typical organisational context, it is rarely that a decision-maker can simply say that he refuses to acknowledge the evidence without raising protests, and possibly doubts as to his

suitability as a decision-maker. For this reason, it is suggested that the typical form of 'secondary elaboration' by decision-makers in these contexts is the attack on the validity of the evidence, We will return to this point in a moment.

In the researcher-client situation, the incorrigibility of the client's propositions is of first importance. Moving on the basis of our analysis of evaluation, we expect two types of 'proposition' to be prone to incorrigibility. We have already suggested that calls for formal evaluation are always based on a preceding informal evaluation. This consists at its simplest of an evaluative scheme, such as:

If x is the case  $\longrightarrow$  good

If not - x is the case  $\longrightarrow$  bad

and a fact statement

x is the case

giving

The state of affairs (being x) is good.

Suppose, for example, we have, as researchers, been called to meet a mythical figure called the Director of Counselling. He begins by telling us he believes the counselling service is 'unsatisfactory'. From this we infer that:

1. he believes the counselling service currently manifests certain characteristics
2. these characteristics are specified in an evaluative scheme held by him as sufficient for the application of the evaluative term 'unsatisfactory'.

Thus we may identify factual beliefs which are incorrigible, in this case the Director's beliefs about the current characteristics of the service. We may also find incorrigibility in the evaluative scheme. For the moment we restrict ourselves to the factual beliefs about the state of the counselling service.

Now the crucial point is that if the client's factual beliefs are incorrigible, research to establish those beliefs as 'knowledge' is, by definition, pointless since an incorrigible belief will not be given up whatever happens. This is, perhaps, so obvious that it hardly needs saying. However, it appears that research usually proceeds in these contexts without any attempt to determine whether the client's beliefs are incorrigible or not. This, we suggest, arises partly because of the researchers' rational-empirical assumption that the client will behave in different ways depending on the results. If the clients beliefs are incorrigible, however, then if the research confirms his belief he will do A, and if it does not confirm it he will refute the research and still do A. Whatever the results of the research, the client still does A.

It is not claimed that clients deliberately commission research with the intention of acting in the way outlined above. On the contrary, they may often be unaware that their beliefs are incorrigible. Nor is it a matter of establishing certain specific propositions as universally incorrigible so that all we must do is see if the client holds them. Rather it is necessary to establish incorrigibility with individual clients in specific contexts.

How can incorrigible propositions be identified? The writer suggests the following tests:

If the client cannot say what information would make him acknowledge his belief as false, the belief can be regarded as incorrigible.

Thus if the Director believes that most counsellors hate counselling, and if he can not suggest any information that would lead him to acknowledge that most do not, his belief is incorrigible. However, it may well be that the client is unaware of certain tests that might applied, so we can widen this to:

If the client cannot say what information would make him acknowledge his belief as false, and will not accept suggested information from the researcher or from any other person, his belief can be regarded as incorrigible.

Thus the researcher may suggest questionnaire data consisting of a 100% response from all counsellors answering the question 'Do you hate counselling?' with 'No'. Notice that the client can quite intelligibly refuse to accept this. He may say, for example, that most counsellors would never admit they hated counselling on an impersonal questionnaire, or because if they did they would be frightened of losing their jobs, and so on with other secondary elaborations. The important point here is that as yet we have not committed resources to research; we have not sent out all the questionnaires only to find that the client doesn't accept the findings. The researcher must, of course, continue this testing procedure until either acceptable tests are found, or the clients' beliefs are finally established as incorrigible. Of course there is no complete finality in this, since tests may be thought of eventually which the client would accept. On the other hand it must be remembered that the client may always be able to invent new secondary elaborations; and again we do not need to necessarily impute bad faith if this happens.

A further problem is that the client may propose or accept a test which cannot be applied in practice. In this case his belief is corrigible in principle but incorrigible in practice. For instance, suppose the Director would not accept questionnaire responses, but would accept interview responses provided there was the same coverage as with the questionnaire. Assuming there are insufficient resources for such interviews, the data which would inform his test cannot be secured. Thus:



If the client cannot say what information would make acknowledge his belief as false, and will not accept suggested information from the researcher or anyone else, where such information must be practically available within the constraints of the situation, the client's belief can be treated as in practice incorrigible.

A procedure of incorrigibility testing of this kind does, however, pose certain problems. The most intractable of these, in the writer's view, is that clients are likely to avoid appearing incorrigible. It is perhaps not respectable to admit that nothing is ever going to convince you that such-and-such is the case.

The extent to which a client can successfully reject evidence, or hypothetical evidence, without being 'unreasonable' is limited. Since we have suggested that the primary grounds for refutation are invalidity arguments, then we suggest that the more the researcher acknowledges threats to validity, the easier it will be for clients to refute findings. To this can be added the qualification that the more esoteric the research subject and methodology the less likely the researcher is to acknowledge threats to validity and the harder it will be for clients to refute findings. By 'esoteric' is meant 'removed from the layman's experience.'

Thus when a patient tells a doctor he has 'pneumonia', if after examining him the doctor declares that he is suffering not from pneumonia but influenza, the patient will have a hard time maintaining

his belief in his pneumonia diagnosis. He could tell himself his doctor is wrong, has mis-read the symptoms, is poorly trained, etc., but it is unlikely. Moreover the doctor is not going to add to his diagnosis that in fact there is considerable uncertainty about whether it is influenza.

In contrast, as we have already said, the social sciences are particularly prone to validity problems and it is regarded as part of the ethic of social research to acknowledge possible causes of invalidity in research reports. These sources (e.g. non-response, experimenter effect, question misinterpretation, lack of control groups, etc) have tended to become common knowledge, partly through the 'popularisation' of the social sciences and the use of similar techniques in a wide range of contexts, such as newspaper surveys and market research. As a result many clients either know, or soon learn, how to ask pertinent validity questions. For this reason, the evaluator, together with other applied social scientists, seems particularly likely to be working with clients whose beliefs are, in practice, incorrigible.

In an earlier chapter, we argued that formal evaluations are most likely to be called for in organisational settings, and hence in all probability by committees. We also suggested that calls for formal evaluations typically arise from dispute, and that to simplify matters we could treat the dispute as being between two parties, whether individuals or groups.

We can now consider the problem of a dispute involving parties holding mutually exclusive incorrigible beliefs, calling the parties A and B. A believes x and B believes not-x. Again it is clear that if both A and B's beliefs are incorrigible, research results cannot resolve the dispute. Instead of confirming the state of affairs concerned as x or not-x, and hence resolving the dispute in favour of A or B, whichever result the research produces either A or B will refuse to accept it. Of course we are assuming that the parties do not realise their beliefs are incorrigible, and are therefore unaware of the futility of requesting research. If the research indicates x state of affairs B introduces validity arguments; if it indicates not-x state of affairs, A introduces validity arguments. Since the researcher will usually be unable to convincingly refute these, and at the same time not be able to admit that they are quite as problematic as the arguer claims, A and B will be no further forward except for having increased the stock of 'ammunition' for their arguments.

It may well be that clients' beliefs are typically not incorrigible. Yet it would certainly be worthwhile for evaluators to devote some time to establishing this before launching into costly empirical investigations.

It is perhaps a truism to say that nobody likes to look a failure or to seem to have taken a bad decision. It is not then surprising to find that many decision-makers are unwilling to accept negative evaluations.<sup>43</sup> The tragic aspect of this is that there is often no reason to attribute a designation of failure or to impute a bad decision

to those who are supposedly 'in control' of social organisations and institutions. If as is frequently the case it is virtually impossible to predict with any degree of accuracy the consequences of the alternative courses which present themselves, it seems hard if things go wrong to attribute blame to the decision-maker.

Certainly if he fails to seek out relevant evidence, or ignores the evidence he finds, or makes an error in his calculations, or chooses in respect of some value which he knows it is not legitimate for him to apply, then he might reasonably be accused of bad decision-making.

But if he chooses on the basis of evidence whose validity is questionable, or on the basis of poorly established relations between the courses of action available and the ends they are intended to achieve, and if he must choose, then if things go wrong how can he be said to have made a bad decision? Yet it seems that even decision-makers themselves believe that they are at fault if things go badly.

How else can the reactions to negative evaluations be explained?

One might perhaps adopt the view put forward by Cleverley that people "are not rational creatures". Thus:

"It is of little use to attempt to change anyone's behaviour by rational argument. It may be possible to convince him that what he believes is wrong, and that what he does is silly. But it will merely change his verbal behaviour. In future he will apologise for what he has done, but he will still do it."<sup>44</sup>

This is, perhaps, too pessimistic. But it does seem to be worth considering to what extent clients for evaluative enquiries are imbued with the 'rational-empirical' philosophy. To help to ensure that evaluative studies are real inputs to the decision-making situation, it may be desirable to engage the client in a modelling exercise in which some consideration is given to what he intends to do in respect of the various ways in which the evaluation may turn out. Ideally, the presentation of 'dummies' of the data which it is intended to collect might be used to identify possible reactions. This is, of course, inherent to the process of setting explicit standards anyway. Such a procedure would at least heighten the clients awareness of the fact that things might not be as he believes they are. It might also help to allay the sense of shock which seems to occur when results turn out to be negative.

REFERENCES

1. Director of IET, Memorandum to the T&CPWG, TC/15/3, 16 June 1970.
2. Ibid.,
3. T&CPWG, 'Report to Senate', S/15/3.
4. Director of IET, 'Research Plans for 1970-71'.
5. IET, '1971 Annual Report', p.8.
6. IET, 'Report on Phase 1b)', July 1971.
7. Ibid., p.6.
8. See IET, 'Successful Counselling Practices', and IET, 'Reports of Visits to Study Centres: Tuition and Counselling Research Project, Phase 2a)', September 1971. See Supplementary Material.
9. Thomas, 'Final Report on the Analysis of Counsellors' Form C4'.
10. A.B. Thomas, 'Report on the 1973 Counsellor Characteristics Study', January 1974, later published as 'Success and Failure in Open University Counselling', See Supplementary Material.
11. IET, 'Report on Phase 1b)', July 1971.  
Thomas, 'An Examination of Drop-out, Wastage and Conservation of Part-time Tutors and Counsellors 1971-72'.  
A.B. Thomas, 'Follow-up of Reasons for Non-contracting of Part-time Staff Completing Contracts in 1971, October 1972.

Thomas, 'Census of 1972 Part-time Staff'.

A.B. Thomas, 'Drop-out and Wastage of Part-time Tutors and Counsellors 1972-73', July 1973.

A.B. Thomas, 'Characteristics of 1973 Part-time Staff', November 1973.

12. Thomas, 'A Summary of Some Research Data on Counselling 1971-72'.

13. As was noted earlier, there seems to be some disillusionment with the experimental approach among evaluators.

14. Open University, Further Particulars for Part-time Tutors, Bletchley, Bucks. 1970, p.2.

15. Ibid., p.3.

16. Open University, Report of Group II at an Open University Conference of Senior Counsellors, 23-25 February 1971.

17. Open University, Report of Group III at an Open University Conference of Senior Counsellors, 23-25 February 1971.

18. Open University, 'Summary of Verbal Reports on Study Centres, Tuition and Counselling', 25 February 1971.

19. Cronhelm, 'Summary Report on the Roles and Functions of Counsellors and the Counselling Service', para. 2.

20. E. Cronhelm, 'Notes on a Brief Analysis of a Small Sample of C3 (1972) Forms', 15 January 1973.

21. Woodley, p.15.

22. Director of Studies, Regional Tutorial Services,  
'Tutor and Counsellor Contact with Students,  
C5 and CT5 Information, January to March 1973',  
4 July 1973.
23. Director of Studies, Regional Tutorial Services, Memorandum  
to the Regional Services Committee, 8 August 1973.

24. 'Early drop-outs' were students who did not 'finally register' with the University. Students taking their first Foundation level course (the lowest level of undergraduate course offered by the University) were 'provisionally registered' for the period January to April at the start of the Open University's academic year. Students 'finally registered' at this time were then committed to the full payment of course fees. Those who did not finally register were thus able to withdraw at a minimum of expense. 'Non-continuers' were those who finally registered but who either did not take course examination, or who took the examination but did not continue to another course in the following year.

The sample of 'early drop-outs' totalled 1338, 66% of whom responded.

The sample of 'non-continuers' totalled 1635, 62% of whom responded.

See McIntosh and Blacklock.

25. Ibid., p.11.
26. Ibid., p.13.
27. Ibid., p.19.
28. Phillips (1973), p.14.



29. Guardian, 'The Softest Cut and the Worst', The Guardian,  
13 August 1975.
30. P. Rossi, 'Boobytraps and Pitfalls in the Evaluation of  
Social Action Programs', in Proceedings of the  
Social Statistics Section of the Annual Meetings  
of the American Statistical Association,  
Washington D.C. 1966, p.129.
31. T. Husen, International Study of Achievement in Mathematics,  
Stockholm 1967.  
  
A. Little and J. Russell, paper read at the UK Reading  
Association Conference, 1971.  
  
J.M. Morris, Standards and Progress in Reading,  
Slough, Bucks. 1966.  
  
R. Davie, 'The Child, The School and the Home', paper read at  
the Annual Meeting of the British Association for the  
Advancement of Science (Education Section),  
3 September 1970.
32. This story is told by G. Cleverley in Managers and Magic,  
Harmondsworth, Middlesex 1973, pp. 73-74.
33. Shipman, p.145.
34. Ibid., p.155.
35. Carter.
36. Weiss, p.126.
37. Carter, pp.118-19.
38. Ibid., p.121.

39. D. Gasking, 'Mathematics and the World', in A. Flew, ed.,  
Logic and Language, (Second Series), Oxford 1959, p.208.
40. Ibid., p.208.
41. See E.E. Evans-Pritchard, Witchcraft, Oracles and Magic  
Among the Azande, Oxford 1937.
42. Ibid., p.330.
43. Berry reports that directors of personnel research departments  
in major American corporations told him that for research  
results to have an altering effect on the organisation they  
had to be dramatic, clear-cut, startling, non-threatening  
and not disagreeable. They must not lay blame, contradict  
decisions or be politically distasteful. D. Berry,  
The Politics of Personnel Research, Ann Arbor, Michigan 1967,  
p.89.
44. Cleverley, pp.195-96.

## CHAPTER TEN

### SUMMARY AND CONCLUSIONS

#### Summary

We began our study by describing the emergence of educational evaluation as a specialism, paying particular attention to developments in the United States. We noted that both evaluation practitioners and their clients seem to be dissatisfied with the current state of educational evaluation. We also noted that a number of different evaluation methodologies have been put forward by practitioners, often using different conceptions of the nature of the evaluation. Commentators on the state of the field have recognised the existence of this diversity of approaches, and practitioners have found it necessary to concern themselves with the question of whether the procedures they advocate might reasonably be described as procedures for carrying out evaluations.

These considerations suggested that a closer examination of the concept might be worthwhile, in the hope that it might be possible to identify some fundamental components of evaluation, an underlying structure which would serve to guide evaluative inquiries in the widest possible range of circumstances.

We also explained that the author worked on an evaluation project at the Open University, concerned with its counselling

service, and that the idea of examining the nature of evaluation more closely arose partly as a result of the problems encountered in the course of this work.

The obvious way to begin was to examine some of the approaches to evaluation made by philosophers in the field known as the Philosophy of Ethics. We therefore summarised several philosophical positions, and noted the implications of these for the conduct of evaluative investigations. We also provided our own possibly novel interpretation of the Naturalistic approach as a means of solving the problem of inconstancy of referents. In the process of our own analysis, we suggested that the further problem of regarding what we later call evaluative standards as analytic or synthetic propositions, could be solved by regarding them as analogous to analytic propositions defined by a restricted community of language users. Evaluative assertions derived from such standards are then verifiable relative to the more or less restricted truth of the standards.

We reached several important conclusions about evaluative terms and statements as a result of our analysis. Of particular importance is that evaluative terms tend to have open criteria of employment, which immediately suggests that any procedure for proxy evaluations, of the kind which professional evaluators are most likely to be called up on to produce, needs to include reference to the need to establish an understanding between researcher and client as to what these criteria are, and to the need for such criteria to be created, where none exist, if the evaluator

is to live up to his title.

We also pointed out that not only are evaluative statements open to requests, or demands, for justification, but more importantly that the criteria for the employment of evaluative terms are themselves open to such questioning. This suggested that evaluations are particularly open to dispute, and we later examined possible sources of disagreement, and ways in which conflicts might be resolved.

In describing a model of evaluation, we envisaged the evaluative process as involving both standard-setting and information getting. Particular attention was paid to the problem of setting standards, and we noted the tendency for significant description sets to become unmanageably large when they consist of variable combinations. We suggested some ways of ordering the set by means of weightings, and then summarised Urmson's discussion of 'grading labels'.

The evaluation of courses of action introduces an extra set of problems, since it is necessary to be able to associate different consequences with each alternative. The possible contributions to be made by workers in fields such as operational research were noted. We then considered possible sources of standards, and emphasised the importance of having justified ones. We also proposed that evaluative schemes could be looked at in terms of their degree of 'sophistication', ranging from single-variable schemes without an evaluative function, to multi-variable schemes with evaluative functions. Then we tried to apply our notion of

evaluative standards to itself by proposing some standards for standards.

In examining evaluative disputes, we said that these could arise either because of disagreements about information, because of disagreements over the evaluative scheme, or because of both. We examined some ways of determining where the sources of evaluative disputes lay, and went on to suggest means of resolving them, or at least of increasing participants' understandings of why they are in conflict with each other. We then argued that calls for evaluative research are most likely to be made under conditions of evaluative dispute, and pointed out that to respond to such calls by undertaking research inquiries is not necessarily an appropriate response.

Turning to the research project concerned with the Open University's counselling service, a description was given of the discussions and arguments that surrounded the establishment of the service, and of the events leading to the setting-up of a committee to propose alternative arrangements. The progress and problems of the Institute of Educational Technology's evaluative research project into the counselling service were detailed. The difficulties were scrutinised with reference to the evaluation model outlined earlier, and other problems arising from the organisational context of the research were noted. Finally we looked at the general problem of the use of research findings by decision-makers, and proposed the use of incorrigibility tests as a means to determining in advance the receptivity of clients to research.

The conclusions which we have drawn from the various parts of the study have usually been stated at appropriate points throughout the text, so they will not be reiterated here. At this point it is only necessary to indicate the way in which we believe we have contributed to the field.

Perhaps the most significant thing we have tried to do is to propose a scheme for the conduct of formal evaluations which was very wide applicability. Many of the approaches put forward by practitioners are principally concerned with the evaluation of instruction, and it seems that the procedures recommended cannot be readily adapted to the evaluation of such phenomena as the Open University counselling service. Schemes which hinge on the use of objectives would not necessarily be helpful if the problem were to establish whether or not it would be desirable to abolish the role of counsellor in the Open University. Although one might be interested in seeing whether counsellors were meeting their objectives, it would also be necessary to examine the likely consequences of abolishing or retaining the role. On the other hand, it could be argued that whether counsellors were meeting their objectives was irrelevant, since the University should not be interested in these objectives in the first place. Under the scheme proposed here, one of the first stages in the research-client relationship is a consideration of what is to be evaluated and why. Instead of responding to a request for research with a research proposal, the researcher is seen as initiating an exploration with the client, or potential client, of the problem which the client believes is amenable to a research solution. A decision to attempt

to measure performance in relation to objectives may or may not result.

The model we have proposed certainly does not claim to be adequate under all circumstances. It has the advantage of being relatively simple in form, although specific evaluative schemes created in its image can be of considerable complexity. We have tried to promote the view that standards require justification if they are to be sensible, although there are obvious limits to the extent to which this can be achieved. The factual aspects of justification may be able to be shown to be false, which can be useful. To repeat an example mentioned earlier, the evaluative scheme that would specify smaller class sizes as better than large ones from the point of view of pupil performance does not appear to be justifiable on the empirical evidence concerning the relation between class size and performance.

In describing the circumstances which surrounded the setting up of the Open University's counselling service, we have certainly been dealing with an unusual situation. However, some of the reactions of the Open University's 'clients' to the research work seem not to have been untypical of those reported by evaluation practitioners working in more conventional contexts. There certainly seems to be some valuable research work to be done in this field. If, as we have suggested, clients usually have expectations and informal evaluations of the phenomena for which formal evaluations are requested, and if they are always, at least in principle, able to find reasons for refuting empirical threats



to their position, then the question arises as to whether research results which run counter to their beliefs and expectations do lead them to modify their initial stances. If research results typically do not have this effect, then we are forced to see the role of research in a new light. Rather than making a contribution to the rational conduct of the enterprise, it becomes a ritual activity not unlike that of the Azande Indians and their chickens.

#### Some Suggestions for the Future

We have already noted many of the problems which surrounded the evaluation research project concerned with the Open University's counselling service. Having considered the nature of evaluation in some depth, and having spent some time describing the events which took place, we will suggest some general guidelines for the conduct of fruitful evaluations over and above those given earlier. We cannot provide any rigid recipe guaranteed to produce results; so much depends on the specific context. The replacement of key client-representatives or researchers by new personnel with different personalities and perspectives, or a change in the financial climate in which the organisation operates, can make a significant difference to the prospects for evaluation. No one can show in advance all the circumstances which might aid or impede an evaluation study.

It is important to reiterate here that we are concerned with evaluation as a process of making value judgements (evaluative statements), and with evaluation research as an activity which supplies data to inform the judgement. It must therefore be

understood that we are dealing here with situations where a client really does want evaluation as we understand it. If he requires, say, a causal analysis or a search for an intervention strategy, the investigator may need to take other considerations into account. Often, of course, the needs of the client will be unclear, so that the investigator has to become deeply involved in the client's problems before he can see what sort of contribution he can make.

○ If there are to be formalized evaluations the most immediate question is 'Evaluation for what?' In everyday life we sometimes make evaluations with no particular purpose in mind other than, perhaps, to communicate our pleasure, or disgust, to others. The resources we use to make such evaluations are, however, fairly inconsequential. Formalized evaluations involving research work tend to be costly by comparison, so it is worth taking some trouble to decide the purpose of any proposed evaluation.

○ In the case of Open University counselling, there seem to be two sorts of purpose which might be served by evaluation.

One is to inform decisions about structural change in relation to counselling. It was this sort of situation which seems really to have been in the minds of many of those who became involved in the research project in its first year. The very existence of the counselling service beyond 1971 was in question. There was indeed mention of its being abolished.

There are two types of structural change that we can envisage. The most radical would consist of eliminating most or all the

objectives of the service, and thus also eliminating the means used to attain, or try to attain, them. This would result, for example, in the abolition of the roles of counsellor and Senior Counsellor, the disappearance of counselling sessions, changes in the central and regional record systems, dissolution of committees concerned with counselling, and so on. It would not necessarily mean losing the services of people employed as counsellors or Senior Counsellors.

A less radical alternative would be to retain some or all of the objectives of the counselling service, whilst changing the methods used to achieve them. The changes in the organisation of the counselling service for 1972 represent a modification of this type, where the declared objectives of the service remained much as they were in 1971.

Clearly decisions concerning major structural changes in the counselling service are the most difficult to inform empirically. If there were dozens of Open Universities, each of which had evolved from a common start, we might be able to find some without counselling services, and so discover what effect this has. Unfortunately, few organisations comparable to the Open University are yet in existence. Failing this, the possibility of experimenting with no counselling service in a single Region, matched with a comparable Region with the full counselling services, remains a useful idea, although the difficulties, both administrative, scientific, and political, are indeed severe. However, even if real-world experimentation is to be ruled out, 'thought-experiments' are still a possibility. We began tracing

out the consequences of a total abolition of counselling above, and much of this can be done without recourse to formal research techniques.

Pursuing these kinds of inquiries produces the factual input of the evaluation. We cannot make a reasoned decision about radical structural changes without some indication of the consequences. Yet we also need to decide whether such changes are desirable. This requires the setting-up and justification of standards which apply to the alternative projections. At this level, this is bound to be difficult, since a radical structural change is likely to have multiple consequences, many of which probably have significance. The procedures suggested in earlier chapters at least enable the task to be approached in a systematic way.

At the opposite pole to considerations of major structural change are changes which leave the basic structure intact. Here we assume the objectives and the major instruments for achieving them are not in question. Instead, attention is focussed on ways of improving the activities which are undertaken within the given structure. For example, the author suggested the use of W.F. Hill's<sup>1</sup> method of group discussion to the Briefing and Training Committee as a way of improving counsellor's group activities, and devised a small research project to obtain feedback on trial sessions.

Improvements can be devised at this level by examining the various activities associated with the counselling service, determining how they are currently carried out, searching for alternative ways of doing them, and then look for ways of making

evaluative discriminations among them.

If we looked at the activity of recruiting counsellors, for example, we might find that applications are stimulated by advertisements in certain newspapers. It would be possible to advertise vacancies in other ways, in other newspapers and journals, by circulars to educational establishments, by television and radio, and so on. The question then is whether any change from existing methods would constitute an improvement, and this would require information on such matters as costs, size of audience reached, type of audience, and so on, and acceptable grounds for ordering the alternatives.

There is, of course, a dilemma over which of these two broad types of decision should be made a focus of formal evaluative inquiries. It could be argued that if the overall structure is insecure, then it is pointless to try to improve the separate parts. This does not seem entirely reasonable, as it is not necessarily a bad thing to try and do the wrong things well so long as it is an open question as to whether they are the wrong things to do. However, it would not seem worthwhile to continually tinker with a system if this means that insufficient resources are left to inform questions about the desirability of retaining it at all.

As formal evaluation usually involves both an investigator and a client, there is always a potential communication problem, made more acute by the inherent variability of meaning in evaluation. We have already seen that communications problems

can have a significant impact on the progress of an evaluative research project, so it seems important to stress the need for some basic understanding between researcher and client. There is an obvious need for the researcher to know what his client expects him to do, and at the same time, for the client to have a reasonable idea of what he can expect from the researcher. It is perhaps difficult for someone who is a specialist in one of the social or behavioural sciences to understand how a client, who may be a layman by comparison, views research. The client may well have unfounded expectations, and be unaware of the implications of proposed activities for his own activities.

Evaluative research, particularly if it is to be carried out in an organisation such as the Open University, often requires the co-operation of many people. Social research in general frequently depends on the voluntary contributions of subjects, but these are usually approached only once or twice for the purposes of a specific project. Moreover they are usually not in a position to present an 'organised resistance' to a particular piece of work. Although it is true that students and to some extent part-time staff tend to come and go at the Open University, the central staff with whom the researchers are likely to have most contact constitute a more stable group. It is therefore likely to be necessary to develop a long-term working relationship with them. This might involve the researchers in trying to communicate what they perceive to be the difficulties of conducting research that will meet the client's needs, and involve the clients trying to give the researchers a better understanding of what they see as their problems.

It should perhaps be remembered that improvements do not necessarily emerge only from evaluative research. Data do not of themselves produce ingenious suggestions. Nor is it the case that proposals for improvement need emanate principally from the centre of the organisation. After all, wherever there are people in the system there is a potential for constructive thought and action. Partly it is a question of encouraging 'improvements thinking', so it seems worthwhile emphasising to counsellors and students, among others, that the University is interested in hearing about ways in which it might be improved.

Evaluation research does not seem to have had a particularly happy history. It would be foolish to conclude that the problems are so formidable that attempts at evaluation should be abandoned, though perhaps equally foolish to underestimate the problems. Hopefully we have been able to make a modest contribution in this thesis to the understanding of these problems, so that future evaluations might be carried out more successfully than has generally been the case in the past.

REFERENCES

1. W.F. Hill, Learning Thru Discussion, Beverley Hills, Calif. 1969. A report on the results of the trials was issued in the autumn of 1972. See D. Armstrong, S. Clennell, D. Gains and A.B. Thomas, 'Learning Through Discussion Pilot Project: Report on Progress of the Project in 1971', November 1972.



APPENDIX I

TUITION AND COUNSELLING RESEARCH PROJECTS INVOLVING THE AUTHOR

1971 Questionnaire to Staff Tutors and Senior Counsellors

Reports prepared by D.G. Hawkridge and A.B. Thomas:

'Tuition and Counselling Research Project: Report on Phase 1a)', March 1971.

'Tuition and Counselling Research Project: Summary Report on Phase 1a)', March 1971 (See Supplementary Material).

1971 Questionnaire to a Sample of Part-time Staff

Report prepared by D.G. Hawkridge and A.B. Thomas:

'Tuition and Counselling Research Project: Report on Phase 1b)', July 1971 (See Supplementary Material)

1971 Study Centre Visits Project

List of successful counselling practices prepared by A.B. Thomas:

'Successful Counselling Practices: Tuition and Counselling Research Project, Phase 2a)', September 1971 (See Supplementary Material).

Series of reports on visits to study centres edited by A.B. Thomas:

'Reports of Visits to Study Centres: Tuition and Counselling Research Project, Phase 2a)', September 1971 (See Supplementary Material).

1971 Monitoring of Tuition and Counselling Report Forms

Reports prepared by A.B. Thomas:

'Analysis of Form C4: Counsellors' Report on Counselling and the Study Centre: Preliminary Report', December 1971 (See Supplementary Material).

'Final Report on the Analysis of Counsellors' Form C4', January 1972 (See Supplementary Material).

1971 Characteristics of Summer School Tutors

Report prepared by A.B. Thomas:

'Summer School Tutors 1971', November 1971.

1971 Turnover of Part-time Staff

Reports prepared by A.B. Thomas:

'Tuition and Counselling Research, Phase 1b), 1972: Examination of Wastage and Conservation of Tutors and Counsellors Between 1971 and 1972: A Brief Report', April 1972.

'Tuition and Counselling Research Phase 1b), 1972: An Examination of Drop-out, Wastage and Conservation of Part-time Tutors and Counsellors 1971-72', April 1972 (See Supplementary Material).

'Follow-up of Reasons for Non-contracting of Part-time Staff Completing Contracts in 1971', October 1972.

1972 Summer School Evaluation

Interviewer on survey of students at summer schools held at Bath, Keele and Norwich.

1972 Census of Part-time Staff

Report prepared by A.B. Thomas:

'Tuition and Counselling Research Phase 1a), 1972:  
Census of 1972 Part-time Staff', November 1972  
(See Supplementary Material).

1972 Learning Through Discussion Project

Report issued by D. Armstrong, S. Clennell, D. Gains and A.B. Thomas:

'Learning Through Discussion Pilot Project: Report on the Progress of the Project in 1971', November 1972.

1973 Census of Part-time Staff

Report prepared by A.B. Thomas:

'Characteristics of 1973 Part-time Staff',  
November 1973.

1973 Turnover of Part-time Staff

Report prepared by A.B. Thomas:

'Drop-out and Wastage of Part-time Tutors and Counsellors 1972-73', July 1973.

1973 Counsellor Characteristics Study

Report prepared by A.B. Thomas:

'Report on the 1973 Counsellor Characteristics Study', January 1974, later published as 'Success and Failure in Open University Counselling', Teaching At a Distance, 1 (November 1974), pp.9-34. (See Supplementary Material)

APPENDIX II

CHRONOLOGICAL SUMMARY OF DEVELOPMENTS IN THE  
TUITION AND COUNSELLING SYSTEM

- 1963 Speech by Harold Wilson on the University of the Air
- 1966 Labour Government issues White Paper
- 1967 Planning Committee established
- 1968 Director of Studies, Local Centres and Tutorial Services appointed
- 1969
- Jan. Planning Committee Report published
- Apr. Director of Studies paper to the Planning Committee:-  
face-to-face tuition rare, counsellor to be main  
student human link with University
- May. Royal Charter granted to the Open University
- Sept. Director of Studies paper to the Vice Chancellor's  
Committee:- clarifies idea of counsellor and view  
on face-to-face tuition
- Oct. First meeting of Tuition & Counselling Project  
Working Group (T & CPWG)
- Nov. Faculties press for face-to-face tuition at T & CPWG  
meeting. Need for proposed level of counselling  
provision questioned
- Dec. Report from T & CPWG to Senate outlines scheme for  
provision of counsellors, correspondence tutors and  
class tutors. Senate approves and suggests priority  
for counselling and correspondence tuition
- 1970
- Mar. Senate approves scheme largely as proposed in December,  
for implementation in 1971

- Apr. Estimates for tuition and counselling provision received by T & CPWG. Dean of Science questions spending on counselling and requests research into the effectiveness of the counselling service  
Institute of Educational Technology established

1971

- Jan. First teaching year begins  
T & CPWG reconstituted as the Tutorial Board
- Mar. Revised system with course tutor and counsellor approved for implementation in 1972

1973

- Mar. Arts Faculty propose a review of tuition and counselling
- Apr. Review Committee on Tuition and Counselling established
- Sept. Review Committee Interim Report favours a model using an academic supervisor and study centre counsellor

1974

- Feb. Review Committee Final Report proposes a model using a tutor-counsellor and a course tutor, which it endorses unanimously.

APPENDIX III

TUITION AND COUNSELLING ARRANGEMENTS

1971 System	Service for students at level*	
	F	Post-F
Correspondence tutor marking and commenting on written work .....	x	*
Class tutor face-to-face tuition .....	x	*
Counsellor general educational and personal advice .....	x	*
*All students in 1971 were on Foundation level courses. Foundation levels courses (F) are 'first year' introductory courses. Post-Foundation level courses (Post-F) are more advanced.		
1972 System	F	Post-F
Course tutor marking and commenting on written work .....	x	x
face-to-face tuition .....	x	x
Counsellor general education and personal advice .....	x	x
subject matter oriented study advice .....	x	

Review Committee's first 'combined model' for 1975 (not adopted)	Service for students at level	
	F	Post-F
Academic supervisor		
marking and commenting		
on written work .....	x	x
face-to-face tuition .....	x	x
general educational		
and personal advice .....	x	x
Study centre counsellor		
specialist personal advice .....	x	x
Review Committee's final model for 1976	F	Post-F
Course tutor		
marking and commenting		
on written work .....		x
face-to-face tuition .....		x
Tutor counsellor		
marking and commenting		
on written work .....	x	
face-to-face tuition .....	x	
general educational		
and personal advice .....	x	x



### 1971 System

Each student would be allocated to a correspondence tutor and a course tutor for each course followed, a different person in each case. He would also be allocated to a counsellor.

In 1971 there were 3564 members of part-time staff employed in 5140 posts as follows:

correspondence tutor	2263
class tutor	1679
counsellor	1198

### 1972 System

Each student would be allocated to a course tutor for each course followed. Each student would also be allocated to a counsellor. In the case of students taking a post-Foundation level courses, the counsellor would often be the same person as in the previous year. Each counsellor would usually have both Foundation and post-Foundation level students allocated to him. For his Foundation level students he was expected to perform the 'defined tutorial role', involving giving subject-matter oriented study advice. The counsellor would normally be familiar with the discipline of his Foundation level students. A counsellor's post-Foundation level students might well be following courses outside the counsellor's sphere, and he was not expected to give subject-matter advice to these students.

In 1972 there were 4351 members of part-time staff employed  
in 4776 posts as follows:

course tutor	3189
counsellor	1587

#### APPENDIX IV

### OBJECTIVES OF TUITION AND COUNSELLING RESEARCH 1971

#### Scope

We are now assuming that five groups of staff should be included in the evaluation of the tuition and counselling system:-

1. Senior counsellors.
2. Counsellors.
3. Staff tutors.
4. Class tutors.
5. Correspondence tutors.

#### Objectives

The evaluation should have the following objectives, all leading to improvement of the system:

- a) Description of tasks expected of and undertaken by T.& C. <sup>\*</sup> staff.
- b) Identification of successful practices.
- c) Identification of recurring problems.
- d) Compilation of checklists of good and poor T. or C. behaviour of each group of staff (1 - 5 above).
- e) Rewriting of job descriptions for 1971 advertisements for T.& C. staff.

\* Tuition and Counselling

- f) Preparation of advisory reports on integration of T.& C. staff into the total instructional system.
- g) Preparation of materials for use during training of T.& C. staff.

#### Activities Monitored

We shall be needing both qualitative and quantitative data, but the categories of data cannot be specified until we have clearer pictures of what the University is expecting some T & C staff to do. We expect that these pictures will be clarified during the next four weeks.

1. Senior Counsellors. Will have to supervise and train counsellors, not clear how. No other activities yet designated. Will be based on Regional Offices. May help with evaluation.
2. Counsellors. Will be based normally on Study Centres. Will meet assigned students roughly twice a month, some students more often, to discuss general problems they have in learning through the OU system. Will supervise Study Centre activities in some fashion. Will seek out students who are signalled as being in difficulties. Will advise students or subsequent courses to be followed. Will assess students' requests to omit summer school.
3. Staff Tutors. Will be based on Regional Offices but will be expected to be at Walton Hall one week per month for contact with appropriate faculty. May play important role in remake process, but this is not clear. Will have to supervise and train tutors, not clear how.

4. Class Tutors. Will conduct tutorial classes (remedial) of up to 20, based on written course materials, using 'notes of guidance'. Will stimulate group discussion. Will submit progress reports on students to counsellors. Will comment on course materials to staff tutors.
5. Correspondence Tutors. Will score and comment upon scripts. Will answer mail or telephone queries from students. Will comment on course materials.

These activities and others to be designated can be summarized in information flow diagrams. The evaluation will seek to meter the flow at various points by collection appropriate data.

#### Data Collection Points

The exact points at which data will be collected cannot be established yet for groups 1 - 4. For correspondence tutors (group 5), there are seven possible data collection points:-

- (i) Assignment scores. These will be on the assessment file for each student in the computer record, and will be readily available. Depending on the assignment, a grade for the whole assignment will be collected, or scores for individual items within the assignment. In both cases, these can be tabulated assignment X tutor.
- (ii) NCR assignment slips. The comments made by the tutors will be on these slips, a copy of which will be at Walton Hall. The slips can be sampled for any tutor or group of tutors and the contents coded into categories.

- (iii) Comments on scripts. If comments other than those in (ii) are made by tutors actually writing on the scripts, these comments can be sampled in the same way as in (ii).
- (iv) Tutors' answers to students' queries. Although no raw data will be available (from listening to telephone conversations or reading letters!) concerning these interchanges, tutors can be asked to complete questionnaires to reflect their view of the kinds of queries made and answer given.
- (v) Comments on course materials. Assuming that these are written, they can be sampled and coded in the same way as in (ii).
- (vi) Comments to staff tutors. Staff tutors can be polled from time to time about comments made to them concerning the course tuition.
- (vii) Comments direct from students. Students too can be polled about their opinions from time to time.

#### Frequency of Data Collection

How often sampling was carried out would depend on what was being sampled. Tutors cannot be asked more than three times in the year to complete special questionnaires for (v) above. For the correspondence tutors, data would probably be collected as often as scripts are marked for (i) above, and 3 - 5 times a year for a tutor marking all assignments for a course for (ii) - (iv) above. Staff tutors and students can be polled three times in a year at most.

#### Sampling Basis

The proportion of potential data that will be sampled will

depend on the accessibility of the data. The computer file allows 100% sampling of (i) above. (ii), (iii), (v) and (vi) must be collected manually, and a 10-20% random sample (say about 500 tutors each time) will be drawn. In the case of (iv), selected (non-random) samples of tutors may be sent special questionnaires, in addition to any random sampling, depending on questions raised within the OU.

#### Data Collection Costs

These must be quite tentative, until the staff activities are clarified. They are the costs involved in (1) paying research staff to read and code from the data documents, (2) compiling, piloting, printing, mailing and retrieving questionnaires, (3) paying for computer tabulations, possibly. We estimate these costs at £2,- 4,000 in 1971.

#### Data Analysis Costs

The amount of analysis required depends on the type and number of questions that must be answered. Correlation of data from sources (i) with data from the other six sources will be expensive, as it must be done on a computer outside the OU.

Again, only a likely range of costs can be given, to cover costs such as (1) paying research staff to examine basic tabulations, formulate, analyses, examine and interpret computer analyses of the data, prepare advisory reports and training materials, (2) paying for computer analysis, including programming, computer input. We estimate these costs at £5, - 10,000 in 1971.

Start-up Costs

Detailed planning of this research must begin in September. IET is prepared to devote some of its limited 1970 research resources to this project if there are assured funds for 1971.

Summary of Costs

Assuming the general objectives and research plan are acceptable to the T & C PWG, the T & C budget may need to be modified to permit an expenditure of up to £14,000 in 1971 on Tuition and Counselling Evaluation.

Dr. D.G. Hawkridge  
3.8.1970

---

This paper was written by the Director of the Institute of Educational Technology as a memorandum to the Tuition and Counselling Project Working Group.



## APPENDIX V

### RESEARCH INTO COUNSELLORS AND COUNSELLING IN 1973

#### 1.0 What do we want to know?

It seems to me that crucial to any decision about how research resources should be allocated is a prior decision about the problems which the application of such resources is designed to solve. Research resources are scarce. They should therefore be allocated to problems which will give a high degree of pay-off relative to expenditure. This implies that we have a suitable analytical grasp of the situation that will enable us to assess in advance the impact of the information we produce as a result of our research efforts. We want to be able to feel reasonably confident that if our results are valid they will have some application to significant decisions.

The first aim of our research should therefore be to diagnose the range of problems which face us. It may be that we are in a position to do this already. Indeed it is one of the functions of the planning group to provide advice as to the kinds of problems to which we might most fruitfully turn our attention. And, of course, there are other groups within the University already actively engaged in tackling problems of various sorts such as the Briefing and Training Committee, the Sub-group on Selection, and so on.

#### 2.0 The Effectiveness of Counsellors

One of the problems which is presumably of some concern to the University is the effectiveness of counsellors. We are

interested in such questions as:

1. What constitutes effecting counselling? How is the performance of counsellors to be judged?
2. Are there any particular characteristics which distinguish more effective from less effective counsellors? Are these characteristics which we can select for, or can they be acquired through briefing or training, or both?
3. What tools, strategies and procedures can be devised which will help counsellors to perform their functions more adequately?
4. What do most counsellors spend most of their time doing and with what results?

Of course, most of these are rather Big Questions, which pose numerous difficulties. And there is probably a certain amount of work already in progress which goes some way to answering these questions. Nevertheless, I think we could make some attempt to pursue at least some of them with whatever resources we have at our disposal.

## 2.1 Descriptive Study of Counsellor Activities

Taking the last question above first, I would like to suggest that a descriptive study be carried out throughout 1973, which would provide a representative description of what counsellors actually do, as opposed to what we think they do or what they ought to be doing. This seems to me to be an essential first step since it would give us a more objective picture of what we are talking about when we consider the counsellor's job. In essence, this would, I

suppose, be a diary study, which would involve selected counsellors in keeping week by week records of their counselling activities. Some information for instance on attendance could be collected from existing records (i.e. the counsellor's attendance register) but information would be sought also on what activities took place and what part the counsellor played in them, e.g. gave a lecture on X, held private consultations with individual students, promoted informal discussion of group study problems. We would be interested not only in activities at the study-centre but in any activity related to the counsellor's work for the O.U., such as dealing with queries at home, reading course material and so on. We would also be asking the counsellor to record any outstanding problem encountered that week, how he attempted to deal with it, and whether he felt able to handle it satisfactorily. This would help to identify in a systematic way recurrent problems which counsellors face and would identify both the most frequent types of problems, and those which counsellors find most difficulty in handling. It might also be possible to reveal in a more detailed way than hitherto the changing pattern of problems encountered throughout the year.

The result of this exercise would be a systematic description of what counsellors actually do in the course of their work, together with a comprehensive dossier of counselling problems, some successfully solved and others not. It may be possible to devise a typology of counsellors from these records possibly on some such dimension as active-passive. Such a typology would be checked out for validity by comparisons with Senior Counsellors' ratings of typed counsellors, and if possible student assessments.

The main difficulties involved (and these should not be underestimated) seem to be:

1. If week by week records are to be kept, this inevitably means a little more work from the counsellor.
2. There is a problem of getting valid information particularly if the respondent thinks that the information he gives may be used in a manner detrimental to his own interests. The best safeguard here is a frank approach to those involved and a guarantee of anonymity.
3. It could be difficult to produce a standardised reporting system giving the varied circumstances under which counsellors work.

However, I believe that most of these problems could be overcome to the extent of making the project worthwhile. The number of counsellors involved would ideally be about 200, but this would bring in 200 records a week or 6,000 over a 30-week year. As our resources will probably not be able to handle these numbers, it may be better to take say 50 counsellors. This would give about 1,500 records in a year.

## 2.2 Counsellor Characteristics

A second area of interest is that of discovering what it is that counsellors do that make them effective, and of trying to relate this to other characteristics which can be identified at the selection stage. A method known as the "critical incidents" technique may be of use to us in the pursuit of this question. I am not, myself,

very familiar with this method, but as Professor Hawkrige has worked at the American Institutes for Research, where the technique was developed, he will be able to give valuable advice on its application. Andrew Northedge is also familiar with the method.

As I understand it, the starting point is an agreed general statement of the general aim of the activity which is of interest (i.e. counselling). From this, persons who are in a position to have seen the persons performing the activity are asked to recall, or specifically watch out for, sets of actions which are conspicuously successful in attaining the general aim. A large number of such incidents are collected and categorised according to some scheme. In this way the critical behaviours which relate to the activity are revealed. These can then be used to give direction to the training activities associated with the activity, and inferences made about the characteristics required to generate the critical behaviours. One of the virtues of the technique is that it helps to eliminate a priori specifications of the relevant characteristics in favour of characteristics which are shown to relate to the activity.

If this aspect of the project were to be pursued, I should like to work closely with Andrew Northedge, whose proposal for a tutor characteristics study would have obvious parallels with this similar study of counsellors.

### 2.3 Procedures for Helping Counsellors to be Effective

Work is already being done to devise ways of helping counsellors to be more effective, such as the provision of

information on group discussion, ways of diagnosing students' difficulties and so on. The specific aid I would like to suggest is the development of a self-evaluation programme for counsellors. This would probably relate to the study of critical incidents which would provide the critical behaviours against which the counsellor would evaluate himself. Not all counsellors will want to evaluate their own activities in a systematic way, but I believe that if some counsellors find it a useful activity to engage in the practice will spread. The benefits of self-evaluation are that:

1. It helps the counsellor to diagnose his own strengths and weaknesses in the job.
2. It serves as a reminder of the job he is supposed to be doing and thus helps to maintain a balanced approach.
3. It helps the counsellor to diagnose and consciously reveal his needs for additional information and guidance.
4. It helps the counsellor to play a creative part in development of his work with the University, and can help him to get more satisfaction from it.
5. It can benefit the students because self-evaluation helps to take explicit account of student needs which might otherwise not be expressed.

Something resembling such a self-evaluation is the C4 form, but there are other strategies which might be tried. For instance, a number of counsellors could meet together regularly to share and diagnose common problems. These could then be passed on to the Senior Counsellor. Counsellors could make tape recordings of their group sessions which will enable them to reflect on what has been

happening afterwards. This can help to sensitise them to features of the situation which might otherwise go unnoticed. Counsellors could also use various forms of evaluation sheet which students would be encouraged to fill out (although the limited experience in the use of Post Meeting Reaction sheets as part of the Learning Through Discussion Project is not encouraging). There are also various self-evaluation checklists with respect to attitudes and perceptions of students, derived from such studies as Ryan's, which some counsellors may find useful.

Really the idea of self-evaluation is mainly concerned with developing an attitude.

#### 2.4 Survey of Attitudes to Counselling

A final project which seems to me worth pursuing, is one which will try to assess the perceptions of counselling which both students and counsellors have. It would try to answer questions such as:

1. How useful students find their counsellor and in what ways.
2. In what ways students think the counselling service could be improved.
3. What counsellors perceive as their main goals.
4. What motives counsellors have for becoming O.U. staff members.
5. What elements of the job are most/least satisfying.
6. ....

Again, some of this information may already be available, though I do not think that there has been a formal survey of student attitudes to counselling.

### 3.0 Conclusion

These are some suggested projects for research into counselling in 1973. I do not imagine that all of them could be carried out. However, perhaps one or two could provide valuable information on counsellors and counselling.

Alan Thomas

November 1972



## APPENDIX VI

### TUITION & COUNSELLING RESEARCH : WHAT SHOULD WE BE DOING?

#### Introduction

It sounds like a good idea to have a project with a general responsibility for looking into the University's tuition and counselling services. There are obviously many questions in this area which it would at least be interesting and in some cases extremely useful to be able to answer. On the other hand, research costs money, takes time, tends to be intrusive and cannot guarantee to come up with answers. So we have problems. How much should we spend on research? Which questions should we tackle first? And what research methods should we use?

The first question is answered for us by the Tutorial Board although its decisions, presumably, may be influenced by recommendations from the research project. Answering the other two questions, however, seems to be extremely difficult, judging by the amount of time spent on them in both informal and formal discussions. This note outlines some of the factors which appear to influence the answering of these questions, in the hope that it will either help to clarify the situation, so that decisions can more easily be reached, or at least explain and perhaps justify the fact that when decisions are taken they are almost inevitably open to justifiable criticism.

#### Research in education generally

It is perhaps helpful to set our research (into relatively new and complex educational system) in the context of educational

research generally. In an article written in 1968 Gage says, "Let us look at where research on teaching has been. As the behavioural sciences go, it has a respectably long history but a regrettably inglorious one..... Positive and significant results were seldom forthcoming and they survived replication even less often. The research yielded many findings that did not make sense, that did not hang together in any meaningful way". He criticises research workers for having looked only at unmanageably broad educational questions and indicates the degree of his reductionist approach by saying "It may well be that a fifteen-minute explanation of a 5-page magazine article is still too large a unit of teaching behaviour to yield valid, lawful knowledge".<sup>1</sup>

These are sobering words, coming from a respected researcher and they indicate the importance of tackling problems at a realistic, manageable level rather than accepting as one's starting point whatever question the "client" asks.

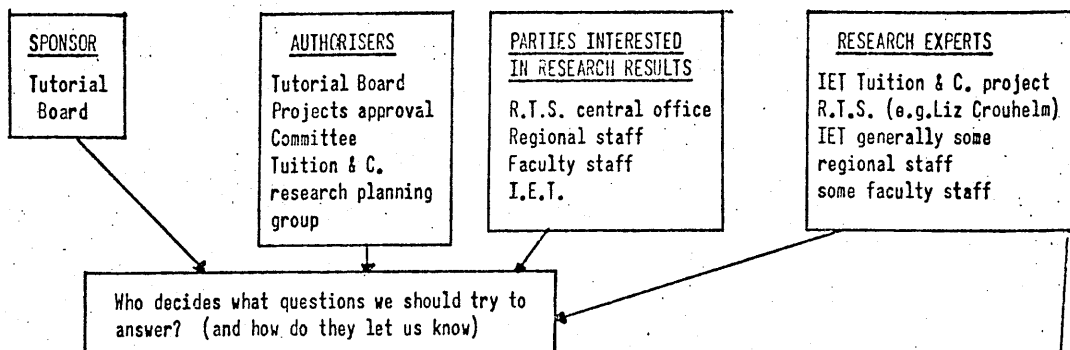
#### Research in the OU

With Gage's rather gloomy caution in mind I turn to the particular problems which confuse the issue of setting up research into tuition and counselling in the OU. On the next page I have tried to schematise these under four main headings.

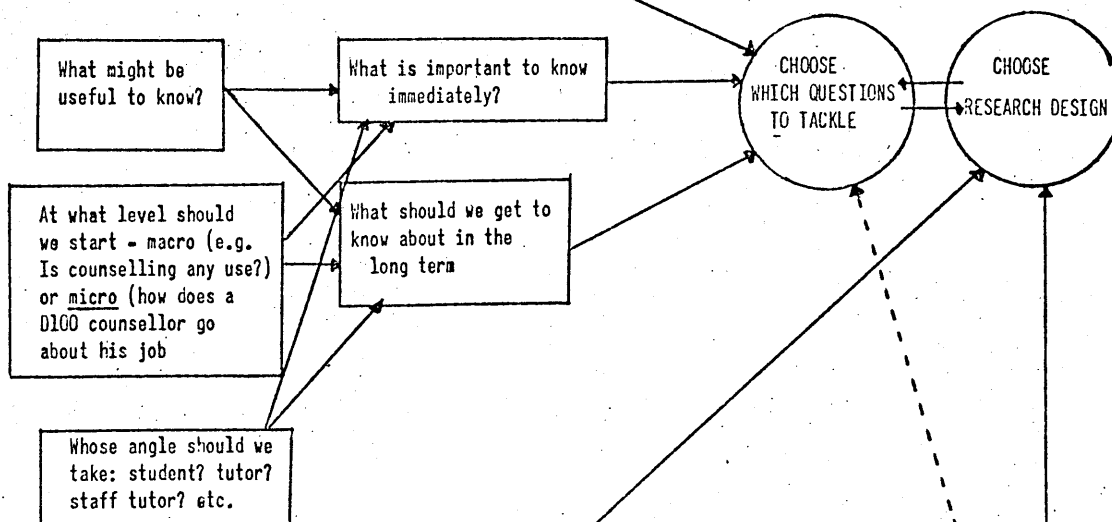
- 
1. N.L. Gage: "An Analytical Approach to Research on Instructional Methods", Journal of Experimental Education Vol.37, 1968, pp. 119-25.

# WHICH STUDY SHALL WE DO? FACTORS INFLUENCING THE DECISION

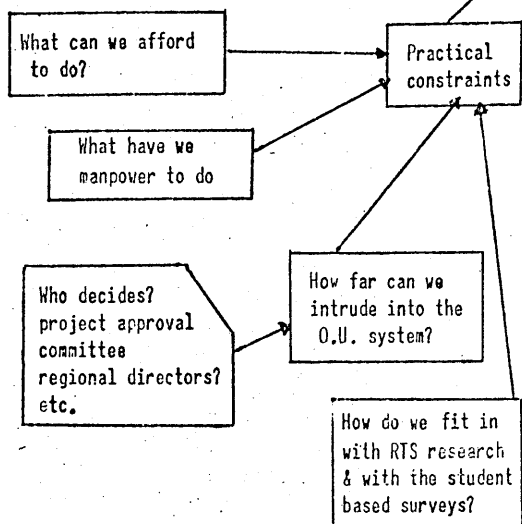
## 1. WHO CALLS THE TUNE?



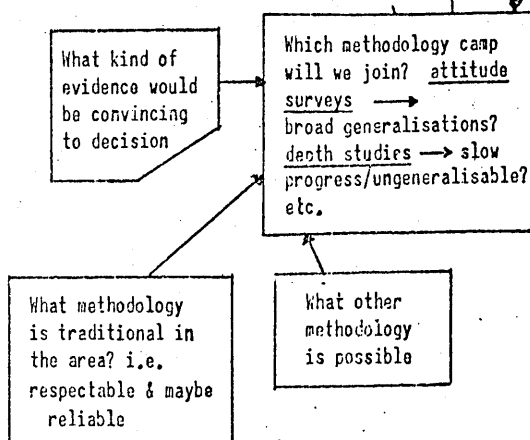
## 2. WHAT DO WE WANT TO KNOW?



## 3. WHAT IS PRACTICALLY FEASIBLE



## 4. WHAT IS METHODOLOGICALLY FEASIBLE



1. Who calls the tune?

- (a) The sponsor: Officially the project is responsible to the Tutorial Board. However this is a large and busy board and its members have varying interests and amounts of research expertise. Consequently we do not and cannot expect to get very detailed direction from it.
- (b) Authorisers: While the Tutorial Board as our sponsor is the ultimate authoriser of our research, we also require the approval of the Projects Approval Committee and (according to current practice) of the Tuition & Counselling Research Planning Group. Other people are also likely to want to have a say in what we undertake including regional staff and faculty staff who might be inconvenienced by our research.
- (c) Parties interested in research findings: Any research project should take its "consumers" into account. These are (potentially at least) almost all the members of OU full-time staff and possibly the part-time staff and students too. In practice we receive quite a number of ideas and suggestions from the Regional Tutorial Service; mainly through the Briefing and Training Committee, since it is particularly active in our area.
- (d) Research experts: The tuition and counselling research project is not the only source of relevant research expertise within the University and consequently our research decisions are open to a fair amount of informed criticism.

An example of the differences in emphasis of the comments, suggestions, etc. we receive from different sources of influence may be taken from two notes received within the last fortnight.

In the first Naomi McIntosh suggests that counsellors have "one overall objective ..... to keep the student in the University".

In the second Liz Cronhelm suggests his "primary function is to assist the educational progress".

These may appear fairly similar and quite compatible suggestions. However, they might suggest very different kinds of research approach.

Another associated problem raised by Naomi's note is that research seems to become very closely tied up with policy-making. At one point she argues the importance of requiring counsellors to return information to Walton Hall on student withdrawals. She includes in her argument that it would force counsellors to keep in touch with students. Assuming that one accepts the desirability of this we might ask how far considerations of "educational practice" should determine our choice of data collection techniques. Obviously such considerations cannot be ignored.

2. What do we want to know?

We start off with a wide range of questions it might be interesting to answer, so where should we begin?

There seems to be a kind of "chicken-egg" cycle which operates whenever somebody says "the important question is ....." Always there seems to be some more fundamental question to answer first, until often we end up coming back round to the original question.

For example:-

"What we need to know is how effective the counselling system is"

"Before we start looking at how effective it is we ought to find out what counsellors are doing, whether it is what they ought to be doing and how far they vary".

"In that case we need to conduct a survey first to find out what staff and students think counsellors ought to be doing".

"We can't afford to spend months findings out what counsellors are or ought to be doing until we have established that counselling is worth what it costs".

This of course leads on to "What do we mean by effectiveness? How do we measure it? How do we evaluate the measured effects?".etc.

One major confusion is whether we should be looking at the broad and correspondingly intractable macro-level problems concerning the value of the tuition and counselling system as a whole or at micro-level problems such as, what advice to give to tutors on the most effective ways of commenting on TMA's. At the macro-level, should we try to evaluate by carefully determining objectives, devising measuring techniques and setting up complex controlled experiments (which might not work)? Or should we be content to concentrate on describing and clarifying the system.

A very real issue is whether or not we should tackle only those problems which will yield usable information quickly. Is it

realistic to expect to arrive at meaningful results quickly? Will a series of quick surveys add up to a useful body of knowledge or will the many assumptions which will have been made be so vulnerable that the results are unconvincing? If we respond to pressures for immediate information will we find we have spent out time unsuccessfully trying to get quick solutions, without having set up a framework of working concepts and usable measures? For example, we need to know within about 6 months to what extent counselling has anything useful to offer. What can we reasonably hope to find out in this space of time? Gage would argue that attempts to answer broad questions are generally fruitless and that you need to start a long way back with a manageable portion of the problem. However, even long term studies cannot guarantee results. Can the OU afford to sponsor research of this kind? And will the problem have resolved themselves in any case by the time the answers arrive?

3. What is practically feasible?

The constraints of money and manpower are obvious and need no comment here. The need to avoid seriously disrupting the system under study is also obvious, although it is not necessarily clear who set the limits, what criteria they use and how they inform us. Another consideration to be borne in mind is the possible overlap between ours and other projects.

4. What is methodologically feasible?

What kinds of answers to our question will we accept? Different research approaches provide different kinds of information. If we

sent a questionnaire to students asking "Does counselling assist you - A lot - A little - Not at all?" would the responses to this present an impressive basis for argument? Would we instead do a more detailed opinion study? Or would we say that students don't know what is best for them anyway and try to get "objective" evidence by comparing drop-out figures for students who have had differing amounts of contact with counsellors? We have to make some kind of assessment of the political context of the study and the plausibility of the kind of findings that a given research method will produce.

It is very likely that an important influence on the method chosen for a particular study is the tradition within that research field and also the past experience of the researcher. Evidence suggests that researchers re-structure problems to fit their accustomed methods rather than choosing the method appropriate for the problem. Possibly this is because their methods are tied up with a whole way of looking at the world. Perhaps it is necessary to become specialised in just one or two approaches to be a successful researcher.

An example of the importance of methodological tradition is the change in attitudes within the University towards the need for item Analysis of CMA's. Because a highly developed methodology of known effectiveness existed, doubts about the desirability of "homogeneity" in tests and of "normalised" treatment of scores have been to a considerable extent set aside for the sake of having some knowledge rather than none. There is no similar methodology for examining the qualities of TMA questions so the case for studying



them is much harder to argue. Similarly, whatever different people may think of the kinds of information produced by attitude surveys, the fact remains that techniques are available for designing, carrying out and analysing such surveys and, again, any knowledge tends to be preferable to none.

The argument is then that, although logically one should establish one's question and then choose one's method to suit it, it may be that in practice successful research is done by the researcher who arrives with his method and decides which questions can be answered by it.

### Conclusions

I would not claim that this is an exhaustive or a balanced account of our problems in deciding where to start. It simply represents an outline of my thinking on the kinds of reasons why research discussions seem to keep revolving in, by now familiar circles. My own view is that we need to take account of Gage's pessimism and tackle problems on a limited scale (e.g. tutors in one faculty within one or two regions) and that we need to take rather more account in our discussions on research what specific questions we are trying to answer, what kinds of answer we are interested in and whether we have methodology available to give that kind of answer. There may also be something to be said for ploughing ahead with a known technique in order simply to get a purchase on the complex problems we are tackling regardless of the many different questions various interested parties may want us

to answer. If our research effort is spread over too many studies, when we are forced by circumstances to use methods which are rather unfamiliar to us and which have not before been used on exactly the same kind of problems, I suspect we will keep coming up with unsatisfactory answers.

Andrew Northedge  
19th January, 1973

## BIBLIOGRAPHY

### Published Works

Ackoff, R.L. and Sasieni, M.W., Fundamentals of Operations Research, New York 1968.

Barker, D., 'Favourable Motorway Meals Survey Hard to Swallow', The Guardian, 19 April 1974.

Berry, D., The Politics of Personnel Research, Ann Arbor, Michigan 1967.

Boalt, G., The Sociology of Social Research, Illinois 1973.

Brehaut, W., 'A Touch of Genius in Us All', Times Higher Education Supplement, 18 January 1974.

Butcher, H.J. and Rudd, E., eds., Contemporary Problems in Higher Education, London 1972.

Carter, R.K., 'Clients' Resistance to Negative Findings and the Latent Conservative Function of Evaluation Studies', American Sociologist, 6 (1971): 118-24.

Cherns, A., Sinclair, R. and Jenkins, W.I., Social Science and Government, London 1972.

Clark, P.A., Action Research and Organizational Change, London 1972.

Cleverley, G., Managers and Magic, Harmondsworth, Middlesex 1973.

Davis, K., Human Society, New York 1966.

Department of Education and Science, The Open University: Report of The Planning Committee to the Secretary of State for Education and Science, London 1969.

- Emmet, E.R., Learning to Philosophize, Harmondsworth, Middlesex 1968.
- Evans-Pritchard, E.E., Witchcraft, Oracles and Magic Among the Azande, Oxford 1937.
- Fleck, A.C. Jnr., 'Evaluation As A Logical Process', Canadian Journal of Public Health, 52 (1961): 185-91.
- Flanagan, J.C., 'The Critical Incident Technique', Psychological Bulletin, 51 (1954): 327-58.
- Flew, A., ed., Logic and Language, (Second Series), Oxford 1959.
- Guttentag, M., 'Models and Methods in Evaluation Research', Journal for the Theory of Social Behaviour, 1 (1971): 75-95.
- Gronlund, N.E., Measurement and Evaluation in Teaching, New York 1968.
- Guardian, 'The Softest Cut and the Worst', The Guardian, 8 August 1975.
- Hammond, P.E., ed., Sociologists at Work, New York 1964.
- Harré, R. and Secord, P.F., The Explanation of Social Behaviour, Oxford 1972.
- Hartman, R.S., The Structure of Value, Illinois 1967.
- Havelock, R.G., A Guide to Innovation in Education, Ann Arbor, Michigan 1970.
- Herzog, E., Some Guide Lines for Evaluative Research, Washington 1959.
- Hill, W.F., Learning Thru Discussion, Beverley Hills, Calif. 1969.

- Husen, T., International Study of Achievement in Mathematics, Stockholm 1967.
- Kaufman, A., The Science of Decision-Making, London 1968.
- Klineberg, O., 'The Problem of Evaluation', International Social Science Bulletin, 7 (1955): 346-52.
- Kourlisky, M., 'An Adversary Model for Educational Evaluation', Evaluation Comment, 4:2 (1973): 3-6.
- Levine, M., 'Scientific Method and the Adversary Model: Some Preliminary Suggestions', Evaluation Comment, 4:2 (1973): 1-3.
- London, J., The Iron Heel, London 1947.
- Martorella, P.H., Concept Learning: Designs for Instruction, London 1972.
- Meehan, E.J., Value Judgement and Social Science: Structures and Processes, Illinois 1969.
- Metcalf, L.E., ed., Values Education: Rationale, Strategies and Procedures, Washington D.C. 1971.
- Morris, J.M., Standards and Progress in Reading, Slough, Bucks. 1966.
- Mulkay, M.J., The Social Process of Innovation: A Study in the Sociology of Science, London 1972.
- National Referendum Campaign, Why You Should Vote No, London 1974.
- , Why You Should Vote Yes, London 1974.
- Open University, Further Particulars for Part-time Tutors, Bletchley, Bucks. 1970.

-----, The Early Development of the Open University: Report of the Vice-Chancellor, January 1969 - December 1970, Bletchley, Bucks. 1972.

Parlett, M. and Hamilton, D., Evaluation as Illumination: A New Approach to the Study of Innovatory Programs, Occasional Paper No. 9, Centre for Research in the Educational Sciences, University of Edinburgh 1972.

Parsons, C.J., Theses and Project Work, London 1973.

Parsons, T., The Structure of Social Action, New York 1937.

Passmore, J., A Hundred Years of Philosophy, Harmondsworth, Middlesex 1968.

Phillips, D.L., Abandoning Method, London 1973.

-----, Knowledge From What? Theories and Methods in Social Research, Chicago 1971.

Popkin, R.H. and Stroll, A., Philosophy, London 1969.

Provus, M.M., Discrepancy Evaluation for Educational Program Improvement and Assessment, Berkeley, Calif. 1971.

Rossi, P., 'Boobytraps and Pitfalls in the Evaluation of Social Action Programs', Proceedings of the Social Statistics Section of the Annual Meetings of the American Statistical Association, Washington D.C. 1966.

Ryans, D.G., Characteristics of Teachers, Their Description, Comparison and Appraisal, Washington D.C. 1960.

- Sanders, J.R. and Cunningham, D.J., 'A Structure for Formative Evaluation in Product Development', Review of Educational Research, 43 (1973): 217-36.
- Sargeaunt, M.J., Operational Research for Management, London 1965.
- Sciara, F.J. and Jantz, R.K., Accountability in American Education, Boston 1972.
- Shipman, M.D., The Limitations of Social Research, London 1972.
- Sills, D.L., ed., International Encyclopaedia of the Social Sciences, 17 vols., 1968.
- Stake, R.E., 'Objectives, Priorities and Other Judgment Data', Review of Educational Research, 40 (1970): 181-212.
- , 'The Countenance of Educational Evaluation', Teachers College Record, 68 (1967): 523-40.
- Stufflebeam, D.L., Foley, W.J., Gephart, W.J., Guba E.G., Hammond, R.L., Merriman, H.O. and Provus, M.M., Educational Evaluation and Decision-Making, Itasca, Illinois 1971.
- Suchman, E.A., Evaluative Research: Principles and Practice in Public Service and Social Action Programs, New York 1967.
- Taylor, P.A., De cort, E and Swinnen, K., 'Standards for Judging Instructional Effectiveness: A Problem in Educational Evaluation', Manitoba Journal of Education, (1970): 5-13.
- Thomas, A.B., 'Success and Failure in Open University Counselling', Teaching At A Distance, 1 (November 1974).
- Thouless, R.H., Straight and Crooked Thinking, London 1953.

Tyler, R.W., ed., Educational Evaluation: New Roles, New Means,  
Chicago 1969.

Tyler, R.W., Gagné, R.M. and Scriven, M., Perspectives of Curriculum  
Evaluation, AERA Monograph Series on Curriculum Evaluation No. 1,  
Chicago 1967.

Unwin, D., ed., Media and Methods: Instructional Technology in Higher  
Education, London 1969.

Urmson, J.O., 'On Grading', Mind, 59 (1950): 145-69.

-----, ed., The Concise Encyclopaedia of Western Philosophy and  
Philosophers, London 1960.

-----, The Emotive Theory of Ethics, London 1968.

Vickers, G., 'Values, Norms and Policies', Policy Sciences, 4:1  
(1973).

Waddington, C.H., Science and Ethics, London 1942.

Weiss, C.H., Evaluation Research: Methods of Assessing Program  
Effectiveness, Englewood Cliffs, New Jersey 1972.

Williams, R.M., American Society, New York 1952,

Wilson, J., Thinking With Concepts, Cambridge 1963.

Wiseman, S. and Pidgeon, D.A., Curriculum Evaluation, Slough,  
Bucks. 1972.

Wittrock, M.C. and Wiley, D.E., The Evaluation of Instruction:  
Issues and Problems, New York 1970.

Wright-Mills, C., The Sociological Imagination, Harmondsworth,  
Middlesex, 1970.



Open University Archive Material

Armstrong, D., Clennell, C., Gains, D. and Thomas, A.B.,

'Learning Through Discussion Pilot Project: Report on Progress of the Project in 1971', November 1971.

Cronhelm, E., 'Notes on a Brief Analysis of a Small Sample of C3 (1972) Forms', 15 January 1973.

-----, 'Some Developments of Counselling', December 1972.

-----, 'Summary Report on the Roles and Functions of Counsellors and the Counselling Service', TB/16/5, January 1973.

Dean of Science Faculty, Memorandum to the Assistant Director of Studies, Regional Tutorial Services, 3 May 1971.

Director of Institute of Educational Technology (IET), Memorandum to the Tuition and Counselling Project Working Group (T&CPWG), TC/15/3, 16 June 1970.

-----, Memorandum to the T&CPWG, 30 October 1970.

-----, 'Research Plans for 1970-71', memorandum to IET staff, 16 July 1970.

Director of Studies, Regional Tutorial Services, 'An Approach to the Question of Personal Tuition', memorandum to the Vice-Chancellor's Committee, 15 September 1969.

-----, Memorandum to the Regional Services Committee, 8 August 1973.

-----, Memorandum to the Chairman of the Review Group on Tuition and Counselling, 12 October 1973.

-----, 'Tutor and Counsellor Contact with Students, C5 and CT5 Information January to March 1973', 4 July 1973.

Grugeon, D., 'Student Counsellors: Their Function and Workload', undated.

Hawkridge, D.G., 'Introduction to Some Aspects of Institute of Educational Technology Research', paper presented at an Open University Conference of Senior Counsellors, 23-25 February 1971.

-----, 'The Role of the Course Tutor in the Open University', paper presented at the Open University Conference on Correspondence Tuition, July 1973.

Institute of Educational Technology, '1971 Annual Report to the Tutorial and Research Boards on the Tuition and Counselling Research Project', undated.

-----, 'Reports of Visits to Study Centres: Tuition and Counselling Research Project Phase 2a)', September 1971.

-----, 'Specific Information on the Institute of Educational Technology', part of a set of further particulars for posts in the Institute, September 1970.

-----, 'Successful Counselling Practices: Tuition and Counselling Research Project Phase 2a)', September 1971.

-----, 'Tuition and Counselling Research Project: Report on Phase 1a)', March 1971.

-----, 'Tuition and Counselling Research Project: Report on Phase 1b)', July 1971.

-----, 'Tuition and Counselling Research Project: Summary Report on Phase 1a)', March 1971.

-----, Minutes of Meetings of the Board, 1970 - 74.

McIntosh, N.E., 'Comments on Alan Thomas' Notes on Research into Counsellors and Counselling in 1973', undated.

McIntosh, N.E. and Blacklock, S., 'Students and Counselling', first draft, undated.

Northedge, R.A., 'Summary of Current Tuition and Counselling Research Projects', September 1972.

-----, 'Tuition and Counselling Research: What Should We Be Doing?', 19 January 1973.

Open University, Report of Group II at an Open University Conference of Senior Counsellors, 23-25 February 1971.

-----, Report of Group III at an Open University Conference of Senior Counsellors, 23-25 February 1971.

-----, 'Summary of Verbal Reports on Study Centres, Tuition and Counselling', 25 February 1971.

Regional Director, East Midlands Region, Letter to the Director of IET, 26 April 1971.

Regional Director, South Region, Letter to the Director of IET, 26 April 1971.

Review Committee on Tuition and Counselling, 'Final Report', February 1974.

-----, 'Interim Report', September 1973.

Science Faculty, Memorandum to the Director of IET, 29 March 1973.

-----, Science Foundation Course Remake Proposal, May 1973.

Social Science Faculty Board, Minutes, 24 February 1971.

Thomas, A.B., 'A Summary of Some Research Data on Counselling 1971-72',  
March 1973.

-----, 'Analysis of Form C4: Counsellors' Report on Counselling and  
the Study Centre: Preliminary Report', December 1971.

-----, 'Characteristics of 1973 Part-time Staff', November 1973.

-----, 'Drop-out and Wastage of Part-time Tutors and Counsellors  
1972-73', July 1973.

-----, Final Report on the Analysis of Counsellors' Form C4',  
January 1972.

-----, 'Follow-up of Reasons for Non-contracting of Part-time Staff  
Completing Contracts in 1971', October 1972.

-----, 'Report on the 1973 Counsellor Characteristics Study',  
January 1974.

-----, 'Research Into Counsellors and Counselling in 1973',  
November 1972.

-----, 'Summer School Tutors 1971', November 1971.

-----, 'Tuition and Counselling Research Phase 1a) 1972:  
Census of 1972 Part-time Staff', November 1972.

-----, 'Tuition and Counselling Research Phase 1b) 1972:  
An Examination of Drop-out, Wastage and Conservation of  
Part-time Tutors and Counsellors 1971-72', April 1972.

-----, 'Tuition and Counselling Research Phase 1b) 1972:  
Examination of Wastage and Conservation on Tutors and  
Counsellors Between 1971 and 1972: A Brief Report', April 1972.

Tuition and Counselling Project Working Group, Minutes and Notes,  
October 1969 - November 1970.

-----, Report, VC/12/1, 8 December 1969.

-----, 'Report of the Tuition and Counselling Project Working Group  
to Senate', S/15/3, 30 June 1970.

Tuition and Counselling Research Planning Group, Notes of Meetings,  
1971-74.

Tutorial Board, Minutes of Meetings, 1971-74.

Vice-Chancellor's Committee, Minutes, 25 September 1969.

Woodley, A., 'An Analysis of a Sample of Completed C3 Forms',  
March 1972.

#### Other Unpublished Works

Crispin, A., 'Part-time Teaching Within the Open University System  
with Particular Reference to the Role of the Part-time Tutor',  
MA (Ed.) dissertation, University of London (1973).

Davie, R., 'The Child, the Home and the School', paper read at the  
Annual Meeting of the British Association for the Advancement  
of Science (Education Section), 3 September 1970.

Hamilton, D., Jenkins, D., King, C., MacDonald, B. and Parlett, M.,  
'Whatever Happened to Daniel Stufflebeam?', pre-publication  
draft, 1975. \*

Little, A. and Russell, J., paper read at the UK Reading Association  
Conference, 1971.

Rowntree, D.R., 'What is Educational Technology?', The Open  
University Institute of Educational Technology Monograph  
No. 1, 1973. \*

Scriven, M., 'Educational Evaluation: Noble Profession, Pedestrian  
Practice', paper presented at an AERA Symposium, February 1971.

\* now published.

THE OPEN UNIVERSITY OF GREAT BRITAIN

SUPPLEMENTARY MATERIAL

TO ACCOMPANY

PROBLEMS OF EVALUATION WITH SPECIAL REFERENCE  
TO THE EVALUATION OF OPEN UNIVERSITY COUNSELLING

A THESIS SUBMITTED  
IN CANDIDACY FOR THE DEGREE OF  
DOCTOR OF PHILOSOPHY

BY

ALAN BERKELEY THOMAS, B.A.

EDUCATIONAL TECHNOLOGY

FEBRUARY 1977

OPEN UNIVERSITY LIBRARY

#### NOTE ON PRESENTATION

The papers and reports included here were originally produced for submission to Open University Committees and members of staff. Although they have been re-typed in order to conform to the Regulations for the submission of theses, the varying forms of headings and tabular presentation have been retained, and references within papers are presented as in the originals.



CONTENTS

	PAGE
1. Tuition and Counselling Research Project: Summary Report on Phase 1a (1971) .....	4
2. Tuition and Counselling Research Project: Report on Phase 1b (1971) .....	10
3. Successful Counselling Practices: Tuition and Counselling Research Project: Phase 2a (1971) .....	20
4. Reports of Visits to Study Centres: Tuition and Counselling Research Project: Phase 2a (1971) .....	26
5. Analysis of Form C4: Counsellors' Report on Counselling and the Study Centre: Preliminary Report (1971) .....	45
6. Final Report on the Analysis of Counsellors' Form C4 (1971) .....	61
7. Tuition and Counselling Research Phase 1b, 1971: An Examination of Drop-out, Wastage and Conservation of Part-time Tutors and Counsellors 1971-72 (1972) ....	73
8. Tuition and Counselling Research Phase 1a, 1972: Census of 1972 Part-time Staff (1972) .....	78
9. Report on the 1973 Counsellor Characteristics Study (1974) .....	83

TUITION AND COUNSELLING RESEARCH PROJECT: Summary Report on Phase 1a

1.0 Purposes

- 1.1 To obtain from Staff tutors and Senior counsellors their job expectations at the start of Open University operations in 1971.
- 1.2 To identify problems (if any) in the use of the University's signalling (message) system for Staff tutors and Senior counsellors as of January 1971.
- 1.3 To list pressing problems (if any) for students, as anticipated by Staff tutors and Senior counsellors in January 1971.
- 1.4 To list pressing problems (if any) experienced or anticipated by Staff tutors and Senior counsellors in January 1971.

2.0 Method

Details of the method are given in the main report (available on request). A 70% return was obtained on the Staff tutors' questionnaire and a 60% return on the one of Senior counsellors.

3.0 Analysis of replies from Staff tutors

- 3.1 Staff tutors' expected use of time:
  - 3.1.1 Time allocations, ranked in descending order, were to work at 1) home desk (or similar place), 2) Regional Office, 3) Walton Hall, and 4) Study Centres.
  - 3.1.2 Travel time was anticipated by 63% to be over 11 hours a month. About 25% said it would be over 21 hours a month.

3.1.3 Types of work at home desk (or similar place) were ranked as follows:

- 1) Follow the foundation course.
- 2) Monitor correspondence tutors' marking.
- 3) Prepare reports and signals.
- 4) Write second-level course units.
- 5) Rewrite foundation course units.

These rankings and those in 3.1.4-3.1.6 below, do not have equal intervals between them, and have been derived from ST replies through an analysis of extremes which examines types of work in terms of the numbers of replies that place each type towards the extremes of the continuum. In other words 'Follow the foundation course' was given priority by more than 'Monitor correspondence tutors' marking', and was placed near the bottom fewer times.

3.1.4 Types of work at Regional Office were ranked as follows:

- 1) Consult with Regional Office staff.
- 2) Read letters from counsellors and class tutors.
- 3) Dictate (secretary or dictaphone) letters.
- 4) Telephone on OU business.
- 5) Scrutinise applications for part-time posts.

3.1.5 Types of work at Walton Hall were ranked as follows:

- 1) Monitor correspondence tutors' marking.
- 2) Attend faculty board meetings.
- 3) Report reactions from Study Centres and Regional Offices.

- 4) Write second-level course units.
- 5) Report errors in the instruction.
- 6) Formulate guidance on marking for correspondence tutors.
- 7) Rewrite foundation course units.

3.1.6 Types of work at Study Centres were ranked as follows:

- 1) Monitor class tutoring.
- 2) Clear up subject matter problems.
- 3) Do some class tutoring.
- 4) Watch/listen to OU broadcasts with students.

3.1.7 Among 20 entries in the 'other response space provided with the questions on overall work patterns were 13 mentions of visiting libraries to do research and to keep up in a subject. The home desk 'other' section elicited 24 responses. Research was mentioned by 13, general administration by 11, and marking correspondence scripts by 9. Among 22 'other' replies under the Regional Office section, general office administration was mentioned by 11.

Among 12 entries in the 'other' response space for Walton Hall, ST course team meetings were mentioned by 14.

Among 20 'other' replies under the Study Centres section, informal meetings and consultations with students were mentioned by 14, with class tutors by 13, and with counsellors by 13.

3.2 Staff tutors' use of signals:

The questions asked under this section were apparently not equally clear to all respondents, but all the replies indicated satisfactory patterns of signalling except those concerned with correspondence tutors marking low.

3.3 Staff tutors' view of students' likely problems:

No single problem outstanding.

3.4 Staff tutors' views of likely problems in tuition and counselling:

Confusion about the counsellors' role, and problems of organizing class tutorials in the 45 minutes available, were mentioned very frequently.

3.5 Staff tutors' view of their own likely problems:

Half the respondents showed concern about being responsible to both Faculty and Region.

4.0 Analysis of replies from Senior counsellors

4.1 Senior counsellors' expected use of time:

4.1.1 Time allocations, ranked in descending order, were to work at 1) Regional or other office, 2) Study Centres, and 3) 'Outside'.

4.1.2 Travel time was anticipated by 13 out of the 14 respondents to be over 11 hours a month. Six thought it would be over 31 hours a month.

4.1.3 Types of work at Regional Office were ranked as follows (using the same method as in 3.1.3-3.1.6 above):

- 1) Scrutinise and act on counsellors' reports.
- 2) Scrutinise and act on counsellors' signals.
- 3) Negotiate re Study Centres.
- 4) Consult with Staff tutors.
- 5) Consult with Regional Director of Assistant Regional Director.

- 6) Reallocate/temporarily replace counsellors.

4.1.4 Types of work at an office other than Regional Office were ranked:

- 1) Negotiate to Study Centres.
- 2) Reallocate/temporarily replace counsellors.
- 3) Scrutinise and act on counsellors' signals.
- 4) Scrutinise and act on counsellors' reports.

The number of respondents is only 8 for this section.

4.1.5 Types of work at Study Centres were ranked:

- 1) Consult with counsellors.
- 2) Monitor counselling sessions.
- 3) Liaison with host institutions.
- 4) Consult with other OU staff.
- 5) Act as replacement counsellor.

4.1.6 The 'other' response space under the Regional Office section yielded only one response, by one person, that was rated at more than 10% of time in the Office; that referred to admissions of students.

Other major tasks mentioned for the office other than Regional Office were mostly to do with record-keeping, following the courses, and answering students' letters.

The outside section yielded immense variety but no pattern. Visiting difficult cases (of students), summer schools, and research were mentioned, for example.

4.2 Senior counsellors' use of signals:

Again the questions were not equally clear to all respondents. There was a good deal of inconsistency

in the replies, but this does not necessarily imply fault in the OU system.

4.3 Senior Counsellors' views of students' likely problems:

No single problem outstanding.

4.4 Senior counsellors' views of likely problems in tuition and counselling:

Not a clear pattern, but the counsellors' role was mentioned most frequently.

4.5 Senior counsellors' views of their own likely problems:

A low response, with no problem outstanding.

5. Recommendations

- 5.1 That comprehensive lists of ST and SC activities be drawn up, based on the categories used in the questionnaires and those mentioned by STs and SCs in their replies, for use by OU personnel responsible for compiling the 'Further Particulars' for ST and SC posts advertised for 1972.
- 5.2 That the attention of the Examinations and Assessment Committee be drawn to the apparent confusion among some STs about procedures in cases where correspondence tutors are marking low.
- 5.3 That the attention of the Tutorial Board should be drawn to the fact that a large proportion of the respondents felt that there was confusion about the counsellors' role in the OU system; and to the frequently expressed concern among STs over being responsible to both Faculty and Region.

TUITION AND COUNSELLING RESEARCH PROJECT: Report on Phase 1b

PART-TIME TUTOR QUESTIONNAIRE

A. AIMS

1. To describe what tasks were being undertaken by part-time tuition and counselling staff.
2. To discover to what extent tuition and counselling staff had received briefing other than printed materials and notes.
3. To discover the background and previous experience of part-time staff.
4. To discover the main sources from which part-time staff had received recruiting information on part-time work with OU.
5. To discover part-time staff conceptions of the students' problems, their own difficulties, and what aspects of the tuition and counselling system, if any, were perceived to be in need of modification.

B. METHOD

The proposed method of data collection was by mail-questionnaire. These were despatched to a sample of 1176 part-time staff on March 19th. Returns were included until April 27th and totalled 942 respondents, representing an 80% return. Total figures in the various tables differ somewhat from the total return of 942 due to various weightings applied to the data in the course of analysis. Percentages based on the national sample should allow valid generalisations to be made about the total population, i.e. all part-time tuition and counselling staff. Percentages on the other tables, i.e. tables for Faculties and in particular Regions, may give less precise estimates due to the smaller numbers upon which the percentages are based. Thus comparisons between these figures and national figures should allow for this.



On Faculty tables figures for Counsellors do not appear, so again this should be borne in mind when comparing Faculty tables with national tables.

In some Faculties some class tutors and correspondence tutors had not started work at the time the survey was undertaken, and many correspondence tutors had been unable to begin work as a result of the postal strike. These groups are thus likely to have made less contribution on tables referring to Jobs x Tasks, Student Problems, Personnel Difficulties and One Thing the OU should Do about Tuition and Counselling.

C. DISTRIBUTION

Copies of the national figures have been circulated to the Faculties, Regions and Regional Tutorial Services, together with Faculty and Regional tables where appropriate. Complete sets of tables are held by Alan Thomas, IET, and David Grugeon, RTS.

D. NATIONAL RESULTS

Table 7: Sex indicates a ratio of 4:1 in favour of males among part-time staff, with a slightly higher proportion of women among people who are only correspondence tutors.

Table 8: Tasks Correspondence script marking had only just begun at the time of the survey. Part-time staff with more than one role were found to be more likely to be following the written, TV and radio components. In general, the table also indicates a lower following for the radio elements of courses than for the TV and written elements.

A high % of counsellors reported that they were dealing with subject matter problems, and many reported that they were reading material relevant to the courses.

A very high % of counsellors reported maintaining contact with Senior Counsellors.

The load of students asking for advice on study methods and self-help has fallen on counsellors rather than on class tutors.

Credit exemption problems were part of the work of only about 20% of the staff to the date of the survey, but about 60% of "Counsellors only" reported dealing with them.

Counsellors reported a very wide diversity of tasks undertaken, compared with class tutors.

Tables 25 and 188: Place of Full-time Work The largest single group of part-time staff worked in universities. The next largest groups were in colleges of education and polytechnics, followed by groups in technical colleges and colleges of further education, and a group in secondary schools. Negligible numbers were employed in industry, local or central government, museums, government research institutes, correspondence colleges, or WEA/extra-mural work. The "Other" category (17% of the sample) included housewives, retired people, post-graduate students, part-timers at places listed above, and the staff of other types of educational institution not listed. Counsellors tended to be drawn less from universities and more from colleges of education and further education than class or correspondence tutors.

Table 42: Teaching Experience in Adult or Higher Education The question was about teaching experience in adult or higher education. As the advertisement for posts specified such experience it was not expected that even a small proportion should have none. The table indicates a small number with 'None', with a concentration among correspondence tutors.

Few said they had any experience of correspondence teaching. Evening classes and day classes were the basis for most of the sample's experience.

Tables 59 and 193: Years in Adult or Higher Education indicate that almost half have less than 5 years experience and in particular correspondence tutors and class tutors are relatively inexperienced.

Tables 76 and 198: Most Recent Degree indicate that about a third of part-time staff obtained a degree after 1966. Counsellors tend to have obtained their last degree at an earlier date, about a quarter of them before 1956.

Table 93: Tutoring Outside Special Subject(s) indicates that almost half of part-time staff had been tutoring partly outside their special subject at the time of the survey. It seems likely that some staff will have interpreted 'special subject' more narrowly than others, however, and of course the policy for employment of generalists varies from Faculty to Faculty.

Tables 110 and 203: Source of Idea of Tutoring/Counselling for OU indicate that about half of part-time staff first received the idea from a newspaper and the rest mainly through work or from a professional journal.

Table 127: Experience of OU Briefing (Other than Notes and Printed Materials) Before 10.1.71 indicates that about one-sixth of part-time staff did not receive such briefing, especially correspondence tutors.

Table 161: Part-time Staff's Conceptions of the Students' Problems indicates that almost half thought that lack of time and the volume of work were major problems for students. To a lesser extent, other problems were felt to be those associated with the course materials and lack of learning skills. Access to books and libraries (including not being able to afford books) and insufficient opportunity to meet with other students, were mentioned as students' problems by only a very few staff.

Table 166: Tutors' and Counsellors' own Difficulties

Overall the majority gave no reply, which may indicate a general absence of problems at the time of the study. Difficulties arising from the course and other materials were mentioned most often and were common to all jobs. About half of those who are only class tutors felt that the degree to which they had contact with students was a problem. Between a quarter and a third of those who are only counsellors were concerned over the nature of their role and difficulties caused by this, and with getting into contact with students.

Table 209: One Thing OU Should Do About Tuition and Counselling System indicates that almost half of the staff gave no reply. The main concern among tutors was for increased or improved face-to-face tutorial contact: suggestions made included more tutorials, longer tutorials, greater use of specialists and smaller numbers in groups. Among some counsellors there was a desire for the merging or combination of the tuition and counselling functions.

E. FACULTY RESULTS

Note: Comparison between tables for the national sample and for each Faculty is difficult on account of there being no Faculty-oriented counsellors. In other words, the Total Sample column on Faculty tables excludes counsellors.

Comments are offered only on those analyses which are considered to be reasonably reliable statistically. Other trends shown in the tables are open to question on account of the size of sample or number of returns.

1. Arts Faculty

Table 9: Tasks indicates that roughly three-quarters of the Arts tutors were following the written, TV and radio components of the course at the time of the survey, and were reading material relevant to the course.

Tables 26 and 189: Place of Full-time Work show that Arts tutors are chiefly drawn from colleges of education, secondary schools and universities.

Tables 60 and 194: Years in Adult or Higher Education Over half the Arts tutors have five years experience or less. The 'Correspondence tutors only' group is least experience of all.

Table 167: Tutors' Own Difficulties Although many tutors reported no difficulties, more than half of those who are 'Class tutors only' felt that the degree to which they had contact with students was a problem.

2. Mathematics Faculty

Table 11: Tasks shows that roughly three-quarters of the Mathematics tutors were following the written and TV components of the course at the time of the survey, but only about half were following the radio component or reading material relevant to the course. Almost all those in 'Class/correspondence tutor' group were following both the written and TV components.

Tables 28 and 191: Place of Full-time Work More than half the Mathematics tutors are drawn from three kinds of institution: universities, polytechnics and colleges of education.

Tables 62 and 196: Years in Adult or Higher Education Nearly half the Mathematics tutors have five years experience or less.

Tables 79 and 201: Most Recent Degree About one-third of Mathematics tutors have obtained their most recent degree since 1966. Less than a quarter obtained their most recent degree before 1956.

Tables 113 and 206: Source of Idea of Tutoring/Counselling for OU

Although most Mathematics tutors heard about the job through the press, about one-quarter heard about it through work.

Table 164: Tutors' Conceptions of the Students' Problems

More than half the Mathematics tutors said that lack of time and volume of work were major problems for students.

3. Science Faculty

Table 12: Tasks Although the figures appear to indicate that rather fewer of the Science tutors were following the written, TV and radio components of the course than was so for other faculties, the differences are not great enough to remove doubt.

Tables 29 and 190: Place of Full-time Work show that Science tutors are drawn chiefly from universities and polytechnics, followed by technical colleges and colleges of education.

Table 40: Teaching Experience in Adult or Higher Education Very few Science tutors have had previous correspondence teaching experience.

Tables 63 and 195: Years in Adult or Higher Education Nearly half the Science tutors have had five years experience or less.

Tables 114 and 205: Source of Idea of Tutoring/Counselling for OU

More than half the Science tutors heard about the job through the newspaper.

Table 165: Tutors' Conceptions of the Students' Problems About 40% of the Science tutors said that lack of time and volume of work were major problems for students.

4. Social Science Faculty

Table 10: Tasks Roughly two-thirds of the Social Science tutors were following the written and TV components of the course, and were reading materials relevant to the course. Slightly fewer were following the radio component.

Tables 27 and 192: Place of Full-time Work show that Social Science tutors are drawn mainly from universities, colleges of education and polytechnics.

Table 44: Teaching Experience in Adult or Higher Education More Social Science tutors have had experience of residential courses than those in other Faculties.

Tables 112 and 207: Source of Idea of Tutoring/Counselling for OU About four-fifths of the Social Science tutors heard about the job either through newspapers, through work or through a professional journal.

Table 163: Tutors' Conceptions of the Students' Problems About 45% of the Social Science tutors said that lack of time and volume of work were major problems for students.

#### F. REGIONAL RESULTS

Regional figures are generally similar to national figures, but comparisons between the two must be treated with caution. All differences indicated below are significant statistically.

Region 1. is generally similar but with a higher proportion receiving the idea of tutoring and counselling through a newspaper and a lower proportion through work. Also a higher proportion did not receive briefing other than notes and printed materials before 10.1.71.

Region 2. is similar but with less people drawn from universities and more who did receive briefing.

Region 3. is similar but with more people whose most recent degree was obtained before 1956 and less people who received the idea for tutoring and counselling from a newspaper.

Region 4. is similar but with less people coordinating services for students, more people drawn from colleges of education, less with 11 years or more experience in adult education or higher teaching and more who did receive briefing.

Region 5. is similar but with more people who did receive briefing.

Region 6. exhibited no significant differences from the national data.

Region 7. is similar but with more people drawn from universities.

Region 8. is similar but with less people answering students written queries and less maintaining contact with staff tutors.

Region 9. is similar with more people drawn from polytechnics.

Region 10. is similar but with less people viewing and listening with students and a greater proportion drawn from universities and none from polytechnics.

Region 11. is similar but with fewer drawn from polytechnics. Polytechnics as such do not, of course, exist in Scotland.

Region 12. exhibited no significant difference from the national data.

#### G. CONCLUSIONS

1. The tuition and counselling system appears to be operating as planned, except that earlier informal reports that counsellors are engaging in subject-matter instruction are confirmed by the data collected in this study.

There seems to be some support for the new pattern planned for 1972, which will "legalise" the Counsellors' subject-matter instruction.



It should be pointed out that it will be difficult to judge whether the new system is of greater benefit to the students.

2. Tutors seem to be more recently qualified and to have had less experience in adult or higher education than the counsellors. Counsellors are less likely to have been drawn from universities. Recruitment has been mainly through the printed word.
3. Almost half the part-time staff considered that lack of time and the volume of work were major problems for students.

#### H. RECOMMENDATIONS

1. That the attention of the appropriate Faculty Boards be drawn to Sections D, E and G.
2. That the attention of the Tutorial Board and Regional Directors be drawn to Sections D, F and G.

THE OPEN UNIVERSITY

SUCCESSFUL COUNSELLING PRACTICES

Tuition and Counselling Research Project: Phase 2a

These lists of practices have been selected from reports made by Staff Tutors and Senior Counsellors as a result of their participation in Phase 2a of the Tuition and Counselling Research Project. The lists are to be submitted to the Tutorial Board Standing Committee on Briefing and Training for consideration as briefing material.

September 1971

Selected and edited in the Institute of Educational Technology

SUCCESSFUL COUNSELLING PRACTICES

- A. The conduct and organisation of counselling evenings
1. Structuring the counselling session.
  2. Getting students to present "seminar papers" in counselling sessions.
  3. Providing time for counsellors to deal with personal problems either before 6.30 pm or after 9.00 pm.
  4. Counsellor directing but not dominating a subject-matter discussion following a radio broadcast.
  5. Setting aside a fixed period for individual counselling.
  6. Teaching the importance of selection and relevance in particular, with regard to study problems.
  7. Counsellors showing students the relationship of the various elements of OU courses.
  8. Counsellors generating critical discussion especially in A 100 and D 100.
  9. Planning the theme of discussions in advance.
  10. Counselling leading discussions without dominating them.
  11. Getting students to make specific contributions to counselling sessions.
  12. Demonstrations by counsellors of those parts of the H.E.K. to be used in current experiments.
  13. When counselling less experienced students, to
    - (a) discuss how to cope with written assignments
    - (b) help students determine and accept a realistic level of aspiration.
  14. Dividing the counselling group into smaller groups on the basis of ability and facility in understanding the course.

15. Regular provision of group sessions and specific invitations to students to attend.
16. Not allowing a few students to dominate discussions.
17. Encouraging students to participate and share problems.
18. Encouraging students to respect standards of scholarship and discouraging attempts to "play the system".
19. Allowing students to have their say - the good practice is to avoid intruding too much and so give the student opportunities to work out his own salvation.
20. Arranging specific times for personal talk with each student early in the session.

B. General counselling practices

1. Counsellors collaborating to provide academic help to students, e.g. sending students with academic problems to the counsellor who specialises in that field.
2. Thorough reading of course-units by all staff.
3. Counsellors must establish personal and professional relationships with students, and course tutors and counsellors must be kept distinct.
4. Self-help groups are valuable but sometimes need the counsellor's "guiding hand".
5. Quick action on the receipt of queries and complaints.
6. Counsellor should manage evening smoothly and create best adult learning surroundings and active social atmosphere possible.
7. Counsellors should keep on good terms with host institution.
8. Observing regulations of host institution and encouraging students to do the same.
9. Counsellors should read, understand and communicate the information and instructions sent to him by his part-time colleagues.

10. Receiving and giving information both to and from students and OU staff.
  11. Counsellors should keep in touch with all students whether attending sessions or not.
  12. Counsellors informing class tutors of academic difficulties before tutorials.
  13. Counsellors arranging group visits to art galleries, theatres, laboratories, etc.
  14. Maintaining friendly contact with class tutors.
  15. Being accessible and genuinely interested in students and treating their individual problems effectively and in confidence.
  16. Promoting the social aspects of the centre by encouraging informal meetings, social evenings, etc.
  17. Keeping accurate and concise records.
  18. Writing a personal letter of introduction to students rather than using the CIL form.
  19. Discussion of plans for counselling sessions at the start of the year, the formulation of a programme and its distribution to students.
  20. Being prepared to give students time on the telephone.
  21. Showing a real interest in assignment gradings and comments and in the general progress of the student.
  22. Frequent demonstration of the Audio-Visual Aids and the provision of a room for use as an Audio-Visual Centre.
- C. The administration of the study centre
1. Counsellor should help manage the computer terminal where provided.
  2. Provision of a table on which students may indicate their time and place of attendance at summer school, encouraging attendance in groups.

3. Signposting rooms in study centre.
4. Arriving early enough to get equipment ready for use.
5. Getting replay devices ready and working.
6. Making course material available for consultation and not allowing it to be taken away.
7. Effective display of notices.
8. Arranging for provision of refreshments when these are not available in host institution.

D. General organisational practices

1. If Senior Counsellor has office in study centre institution it helps administration and liaison.
2. Counsellor's forms, especially C2, help communication.
3. Meetings of part-time staff within and between centres are useful, and travelling expenses should be allowed for this.
4. Access to college bar for social relaxation at the end of the evening.
5. Conducting a day school.
6. Meeting new applicants at study centres for advance view of study centre activities.
7. Duplicated explanatory material on a study centre and its activities, to be provided by the Region, for counsellors to send with their first letter to students.
8. Initial meetings should be carefully planned and give some introduction to the courses as well as explaining the procedure.
9. To impress on counsellors that 6.30 pm to 9.00 pm at a centre is not the limit of their commitment to the OU.
10. Periodic group discussion meetings with Senior Counsellors to keep up to date with information.

11. For new counsellors:

- (a) Fullest possible information in duplicated or printed form about the OU, the chain of communication, etc.
- (b) Description of the centres, duties, work and administration for which the counsellor is responsible.
- (c) If appointed before the end of a session, an invitation to visit a centre to meet existing counsellors and students.

12. Reminder cards, e.g. "Your next D 100 tutorial is on ..... at .....", to be distributed by counsellors to students as appropriate.

THE OPEN UNIVERSITY

REPORTS OF VISITS TO STUDY CENTRES

Tuition and Counselling Research Project: Phase 2a

Here is a selection of edited reports of visits made to study centres by Senior Counsellors and Staff Tutors contributing to Phase 2a of the Tuition and Counselling Research Project. All identification of students, tutors, counsellors, study centres and authors has been removed. The reports cannot give a wholly representative picture of the activities and views in the centres, but offer much interest to Open University Staff.

September 1971

Selected and edited in the Institute of Educational Technology

N.B. For reasons of space, only reports on two of the six study centre visits included in the original report are reproduced here.



## STUDY CENTRE A

### 1. Impressions of Study Centre A

The students are primarily interested in obtaining credits in their foundation courses as a first step towards an OU degree. The facilities of the study centre are adequate but not outstanding, with classroom furniture and modern F.E. college atmosphere. Live and recorded broadcasts are not in great demand at the centre, viewing being mainly at home. Coffee is not available, nor is there any great demand for its provision. The college library is open to the OU students and is generally used either by those who wish to study in pleasant surroundings or by small groups who may wish to study together on nights when their counsellor is not present. The Study Centre Consultative Committee is functioning, but its students tend to regard it at best as a potential action route rather than as an important continuing feature of their community life. In cases where the correspondence tutor had another function and was therefore meeting the students whose scripts he marked, this was regarded as decidedly advantageous. Part-time staff were unanimous in regarding their relative isolation from staff tutors as disadvantageous. The group as a whole seemed conscientious and industrious, and likely to derive increasing benefit from the developing OU system.

### 2. Case Study of Work of Counsellor

The counsellor has made himself fully acquainted with the facilities of the study centre and has made this information

available to all students. As he arrives on the counselling evening he checks if any messages have been left for him and ensures that current notices are on display. He checks that TV and radio are switched on at appropriate channels and that the latest TV and radio cassettes are available.

As the students arrive he establishes what each wishes to do that evening, giving particular attention to any who can rarely attend. Those who wish to view he directs accordingly. For those who wish individual interviews he arranges suitable times, and may proceed with some of these interviews during the TV broadcast.

Immediately the broadcasts end, the counsellor will encourage a group of students to discuss the programmes or relevant units, and will if necessary suggest appropriate topics for group discussion or group study. Where necessary he will suggest that furniture be rearranged to ease establishment of a group working atmosphere.

By 7.30 pm most of the students who intend to visit the study centre will have arrived. The counsellor will make any necessary announcements arising from information he has received recently, and will develop any questions of general interest raised earlier by students.

According to his own knowledge of the progress of the students present, the counsellor will then suggest a return to work in one or several groups. Depending on his own academic

background, on his knowledge of recent course units and of topics of immediate interest to the students, the counsellor may take a more or less active part in some of the discussions. His role at this stage is to initiate group study, not necessarily to lead the group. He makes himself available to each group to answer such questions as he can and to suggest methods of dealing with student difficulties.

He sorts the questions he should himself refer to a senior counsellor from those he urges the students to take up individually with correspondence tutor or class tutor, or as a group through their own Study Centre Consultative Committee. He is ever alert to signs of ineffectiveness in the system and ensures that these are notified to his full-time colleagues as appropriate.

Depending on how self-sustaining the discussions are, the counsellor may wish to return to interviews with individual students. According to the information he has on the progress of individual students, as indicated by assignment results, he may wish to approach a student himself and suggest that some discussion of his difficulties might be useful.

As the evening draws to a close at 9.0 pm the counsellor ensures that students are acquainted with the dates of the next meeting and reminds them of any looming deadlines. He sees to it that the room is left as the host institution would wish, and he checks that he has a note of any points he must raise

with colleagues or with students who have not appeared at the study centre.

3. Case Study of Work of Tutor

The class tutor knows what course units and aspects of course units are to be covered in each tutorial class. Moreover, he has satisfied himself that each and every student likely to attend his class has had this information either verbally from the tutor himself or from the counsellor at an earlier meeting, or by letter or telephone. Thus he appreciates the extent of the background he can assume the students to have and he knows precisely the nomenclature and definitions used in the course material. He may well wish of his own knowledge to expand this material, and to recommend and encourage background readings but he does not raise unnecessary alarm or confusion in the minds of the students. The students having had adequate warning to study the course material and to discuss any difficulties among themselves and with the counsellor, when the tutorial class is held optimum use is made of the limited time available.

The class tutor may have learned either directly from the students or indirectly from the counsellor or correspondence tutor of specific difficulties which have been occurring commonly with the material of the appropriate sections of his subject. If so he arranged his talk to give due prominence to these points. Either he gives a formal lecture for thirty minutes and then asks for specific questions, or he presents his

material in such a way as to encourage questions throughout the forty-five minutes of the class. To a very large extent his choice of technique will depend on his own personality and his knowledge of the individuals in his class. He presents his material not as an isolated lecture on a single topic, but as an explanatory talk on a particular facet of an integrated foundation course.

As he ends his tutorial class, the class tutor will remind the students of the date and subject-matter of the next tutorial meeting and will suggest that they should give time to the preparation of the relevant material. He will welcome any suggestions as to changes of presentation which might be incorporated in his next tutorial class and, if he considers the group as a whole might benefit, may experiment at the first convenient opportunity.

#### STUDY CENTRE B

##### 1. The Work of the Counsellors

###### Counsellor A

Dr. ---, a counsellor in Science, is also a tutor in Chemistry, a member of the OU Regional Consultative Committee, and Liaison Officer with the College where the study centre is situated. In the College, his day-time function is that of Head of the Department of Science, and he forms part of a group of the College staff which includes the Principal and about

six other teaching staff who take a very keen interest in OU work. He evidently plays a very significant part in organising a splendid study centre. His personality is vital and extrovert, communicating his enthusiasm very readily to others. He claims his counselling and tutorial sessions take the form of "just talk" - he adds that he probably talks too much, but we found him also very ready to stop talking and listen attentively, when asked. None of his students unfortunately was present during the visit, with whom we could have further discussed the counselling and tutoring, but we received an impression of a very thoughtful approach to OU counselling on the part of Dr. ---.

He considers the personality of the Counsellor and his relationship with the students to be of prime importance, and is sensitive to the good and ill effects his own personality may have on different types of student. Being readily on friendly first-name terms he is nevertheless aware that this appearance of being easy to know may lead to the facile assumption that he knows the students well enough to give advice whereas the relationship should be more cautiously built up. The first three months, he felt, were critical in this respect, and only after this period did he feel able to offer them advice. If a student was finding the work too onerous, for example, in the early stages, he would not press too hard the possibilities of extra help, being unable to predict the result if hopes thus raised were proved to be false.

He is not a good counsellor, he feels, with certain types of student: those who are already well qualified, and coping well with the OU work for instance, might tempt him to "compete" rather than help, and a sense of rivalry might bedevil the whole relationship. He wonders if such students might rather "spoil" group counselling, but nevertheless would not consider "streaming" his students. He would work far better, he says, with students from a more "deprived" kind of background, who have little genuine self-confidence: here, he could develop an empathy and try to influence the students' attitudes and self-expectation. From his experience teaching at different levels, he typifies such students as having a paradoxical attitude to study: on one hand, they are already conditioned to anticipate failure, while on the other, they expect they ought to be able to grasp and to remember the total amount of material put before them, and to do so immediately. He would attack this attitude, "make the student feel like a rebellious 17-year-old", and try to work out with him what a suitable "target" in assignment grades might be, bearing in mind that the same standards of marking will apply to other students far more privileged than himself. His progress should be measured against his own previous performances, for the first few months, rather than against that of the other students. Only when they had reached the stage of actually over-compensating, would he attempt, very gradually, to raise this student's own level of aspiration.

Additional help in annotation, abstracting the important parts of a unit, etc., can of course be more easily given, but he feels strongly that a personal influence can be a most potent factor in learning (by "osmosis"), and that ideally a student should be offered a choice of Counsellor (an idea unworkable in practice, he admits). He realises that many students might prefer a quieter, more impersonal approach than his own.

All students need someone to watch them learning, and to this end he sees the combination of the roles of Tutor and Counsellor as essential.

#### Counsellor B

Mr. --- holds a degree in Economics and Geography and is head of the Department of Business Studies at the College. (No fewer than seven of the College Staff hold part-time posts with the OU and the College Principal is himself an enthusiastic OU supporter. It is hardly surprising to note that the OU is exceptionally well provided for in the way of facilities).

Mr. --- is also a class tutor, one of three Social Science class tutors at the Centre, but he has not yet acted in this capacity. Being a class tutor, he obtains copies of the course units at home, which obviates problems which have to be faced by many other Counsellors.

Mr. --- explained that the area did not receive BBC 2, but that the video/cassettes were played on the same evening as viewers in more fortunate areas would receive the programmes (i.e. Mondays).



Counselling normally took place on the same evening.  
Average attendance was 12-14 out of an original 23.

The area did receive VHF Radio and the College was equipped with a Hacker set. The broadcast was received on Wednesday evenings, but was not generally well attended. However, if any student should attend, counselling would be held if requested. On this particular evening, one student attended for personal counselling.

On Mondays, Mr. --- said he would normally begin by making any relevant announcements; then the students would watch the TV programme; afterwards he would throw the floor open to the students, who would bring up points from the programme or units, and a good discussion would invariably ensue. The discussion would generally be centred round the programme.

Wednesday's audiences were generally much smaller and counselling was more frequently individual, as it was on this particular evening.

Asked if the counselling job had turned out as he had anticipated, Mr. --- said that he found himself doing much more work than he would have expected (not that he necessarily minded this). Officially, he came in to the centre on Wednesday nights on College duty, but he invariably spent it on OU business. In the course of his work as Counsellor, Mr. --- became virtually a tutor. The whole counselling issue had been widely debated before the OU became operational. The Regional Director

insisted that the Counsellor should fulfil a non-academic role, but Counsellors themselves knew that they might be forced into becoming tutors. In any case, if a Counsellor were interested in the subjects taught he would want to become academically involved in them.

Asked why he took up his appointment, he said that he had been interested in the whole concept of the Open University while it was still on the drawing boards. He wanted to develop modular courses throughout Further Education, as most courses there were not, as yet, modular.

He was also keen to see the various media developed for educational purposes, and thought that the OU promised a real breakthrough.

He was also keen to see the College become an OU Centre!

There was little comparison between the OU post and his full-time post, he said. The latter, although academic in principle, was in practice largely administrative.

Asked what aspects of his previous experience had been most useful in working for the OU he said that he had been counselling students (especially older students) for many years. He had taken a leading part in designing and developing GCE level and other courses, and had conducted a highly successful series on "Counselling" at the College, to which various types of students came from near and far.

He had not given up any other academic activities to work for the OU, though he had given up some domestic activities!

He said that he had already applied to do both Counselling and Course Tuition for next year.

Asked what he liked about his post, he said he liked the feeling that he was playing a useful role in helping mature people who couldn't see the wood for the trees, who seemed to be so obsessed with the details of the course and who needed to see it in a broader perspective.

His one real dislike was the amount of work (including paperwork) involved. Asked about successful practices in counselling, Mr. --- said that the idea of having people together in front of the television set and discussing things after the programme with them was probably one of the most helpful, the programme providing a very good focus for discussion.

Asked about his students, he said that motivation was high, as was attendance (12-14) out of a nominal 23. As far as he knew, no students had as yet definitely dropped out.

On his relations with the Class Tutor (the two other than himself), Mr. --- described these as excellent. He saw them on tuition evenings, and was regularly in touch with them by telephone in between sessions. There was a continual flow of information in both directions.

Asked how he decided what should be discussed at counselling sessions, he said that he left this up to the students themselves. Basically, every matter raised was raised spontaneously. He did, however, try to "guide" the discussion into certain channels. He could, when conducting a proper Tuition session, prepare a structured lecture and deliver it.

Relations among the students were easily handled, so the group were a fairly happy and harmonious entity. Although some were naturally more vocal than others, nobody dominated the group, nor were there any particularly shy students. If there were, he would gently question them to try and obtain some participation. No student had dropped out, so far as he was aware, but weak students or those with discouraging marks could come and see him individually.

## 2. A Counselling Session

This was certainly not typical of a Monday counselling as related above, but was more typical of the smaller Wednesday sessions.

A Mr. ---, a Social Science student from a remote village, came in with a complaint about the "F" mark he had received for assignment D200 04 (Economics). All his other marks had been "C's" and he felt that this particular essay had been unfairly marked. He had put considerable work and effort into the assignment (he admittedly chose a very difficult topic) and regarded the grade he obtained as almost an insult. He felt he should have received a "C" or a "D".

The Counsellor, who is an economist by training, read the essay thoroughly. He said that the essay was unquestionably well-written, knowledgeable, concise and well documented. However, he was forced to agree with the correspondence tutor that Mr. --- had not strictly adhered to the question set in the course material. He had failed, not on his knowledge of economics and of all the concepts and principles involved, but on his understanding of the question asked. However, he confessed that the question was itself badly put and ought to be rephrased. In view of this, and considering the general difficulty of the question, the Counsellor said that he would raise the question of the mark with the correspondence tutor in Social Science for the region.

Mr. --- seemed happy with the proposal.

This session lasted about half-an-hour, and seemed to have been very competently and satisfactorily handled.

### 3. Interview with Students

#### Student A

The student in question was Mr. ---, a 40 year old primary school headmaster, married with four children.

I began by asking Mr. --- what was his object in taking an OU degree. He replied that he had always intended to do a University degree. However, his secondary education had ended in war-time, and he went straight to the Army. On his return

from the Force, he was determined to get into teaching as quickly as possible. He therefore turned down the offer of a University place and took a two year course at T.T.C. After leaving college, he taught both in primary and boarding schools, and had been a headmaster since 1960. He now felt he had the time to take the degree he had always planned to do. Asked for his impressions of the study centre, he said he regarded it as vital, owing to the absence of BBC 2 in the area. The fact that students virtually had to attend made it far more interesting, successful and lively place. He was almost sorry that BBC 2 was coming to the area soon, as he felt that not so many students would come to the centre after that. Asked how frequently he attended, Mr. --- said that he came at least once and frequently twice per week.

I then asked him what he normally did at the study centre. He said that at Monday evening counselling, a discussion would be held on the subject of that evening's programme and the correspondence unit, guided by the Counsellor. Wednesday evening counselling was usually smaller and more personal in nature. He occasionally watched the Arts TV programme on Wednesdays also, and observed Arts counselling.

On tutorial Mondays, there was structured specialist tuition. There were three class tutors, including Mr. --- the Counsellor, though he had not as yet performed in his capacity of class tutor. On a couple of occasions, outside lecturers would be invited to give highly specialist lectures.

A philosophy lecturer from outside had given an evening's lecture on Hobbes. (Perhaps more use could be made of this idea). The quality of tutorials varied; in some cases the tutor merely reinforced what had already been put across in the Units. Others gave their own thoughts on the subject, which was far more useful and valuable.

Asked what he particularly liked about the centre, Mr. --- had little to add to what he had already said. He had no particular dislikes either.

Asked what he thought was successful or unsuccessful about the centre, he said that there was no lack of success. His only criticism was the variation in the quality of the tutorials, already noted above.

Asked about his motivation to attend the Centre, he said that this was generally very strong at the moment in view of the non-availability of BBC2 if the present good discussions continued. Bad tuition might tend to keep people at home, though.

Asked about the course in general, Mr. --- said that it was excellent. However, he noted discrepancies in the standards among the various disciplines, and this could not be put down to the mere fact that students were more familiar with some disciplines at the outset. He felt that the Economics and Geography parts of the course were inferior and more difficult to follow.

Asked about the quality of the programme, he said that the TV programmes were, on the whole, very good, though he again noted discrepancies. The programme on Unit 18 featuring the actor who took part in the survey at Paddington Station was particularly to be commended, but some of the economics block were not so good, especially that on Unit 10, which could have been put to much better use. This wasted a great deal of time trying to explain what graphs were. As it seemed that the authors were assuming so much, they could perhaps have assumed that students were generally aware of what a graph was.

He did not get so much from the Radio Programmes unless he listened to them more than once. He therefore took tape recordings of these broadcasts so that he could listen to them in a more receptive mood.

Asked what he felt about the part-time staff, Mr. --- had little to say about the tutors other than what he had already noted. He again stressed the variation in the quality of the tutorials.

The Counsellor was a painstaking and understanding advisor who also had a good grip of the course material. Mr. --- was very pleased with the "face to face" aspect of the OU generally. This was what was generally lacking in correspondence courses.

Generally speaking, Mr. --- was quite happy with TMA questions and with the grades he received. He was not unduly unhappy with CMAs, but felt that model answers ought to be provided.



Student B

The student, following course A 100 and B 100, makes a half-hour journey by road twice every week, to visit the study centre. The immediate reason is that BBC 2 reception being as yet unavailable in the area, he must see the programmes on the technicolour projector, although he would come in any case "in case he missed anything". Tutorials and counselling are useful only rarely, and he places little value on the radio programme, so that the real attraction for him is the company of other students. All his friends at residential universities told him this is what he would miss at the OU, and he has come to agree with them.

He seems to be an extremely capable young man, who left school without university entrance qualifications through being "too lazy", and more concerned with the school Debating Society, the canoeing club and so on - all an excellent preparation, he says, for the OU course. He gives the impression of being more than capable of solving his own practical, domestic and administrative problems, and hence sees little value in the concept of "Counsellor", although he would admit that when the latter gives an impromptu "tutorial" on his own subject - Economics - he finds it very worth while.

Otherwise, he finds the Counsellor's role difficult to define, and also that of the Tutor: the time available is inadequate, the groups too large, and the subject-matter out

of phase with both the correspondence texts and the assignments. His daytime occupation is teaching Art and History at a Friends' School, qualified by a general Teacher's Certificate; career improvement in his main motive in taking an OU course. So far, he is disappointed by the level of work, which is boring and mechanical until he comes to a block which is more difficult, and so far, only Economics had been really "stretching" for him. His assignment grades average "B", and he claims to spend only about 10-15 hours per week on "real study", as distinct from discussion, obtaining books and so on, and on the face of it, much of what he says about himself and the study centre may be true. Yet, on the evening of the visit, there was a splendid tutorial, followed by a thoughtful yet lively discussion, held in ideal surroundings - the staff coffee room of the College. This tutorial, in Art History, was also attended by the group's tutor in Literature and Music who had come along for sheer interest, accompanied by her husband who had come for the same reason (although he would not take part in the discussion). This tutor acts also at another study centre and declares the students, staff and accommodation are very much preferable at this Centre thus bearing out our own impression that the enthusiasm of the OU Staff, the whole-hearted co-operation from the college, the lively and intelligent response from the students all combined to make Wednesday evenings at this study centre one of the most promising in our experience.

ANALYSIS OF FORM C4

COUNSELLORS' REPORT ON COUNSELLING AND THE STUDY CENTRE

Preliminary Report

1.0 Introduction

1.1 At a meeting of the Tuition and Counselling Research Planning Group held on October 5th, it was suggested by a member of Regional Staff that a rapid national summary of the contents of C4 forms could be had by providing Senior Counsellors with summary sheets. The idea was taken up and coding sheets produced to cover most of the questions on the form. These were despatched to Senior Counsellors with explanatory notes. For questions where categories could not be devised Senior Counsellors were invited to select pertinent comments from their Counsellors' C4s and to make any other additional comments they felt to be important. It will only be possible to include a few of these here; a fuller set will accompany the final report.

1.2 This preliminary report is based on information received to date from 13 Senior Counsellors, representing responses from 457 counsellors. All Regions are represented except 5, 6, 10 and 12.

2.0 Results

2.1 Question 1 a)

Stated: 'Apart from regular meetings at the study centre, I

came to know my students in the following way'. The form gave five alternatives, any or all of which could be ticked. The results were:

		%
1. Telephone	378	82.7
2. CIL introductory letter	237	51.9
3. Other written contact	228	49.9
4. Induction meeting prior to January 1971	222	48.6
5. Home visit	115	25.2

This would seem to indicate a widespread use of telephone contact between Counsellors and students. One Senior Counsellor pointed out that home visits are normally made at the Counsellor's own expense.

## 2.2 Question 1 b)

Stated: 'Apart from regular meetings at the study centre, the following have been the most valuable means of contact (to me and/or the students) and for the following reasons'.

Answers here fell into largely the same categories as for Question 1 a). The results were:

		%
1. Telephone	266	58.2
2. Written contact	136	29.8
3. Other group meetings	77	16.9
4. Other means	63	13.8
5. No answer/meetings at study centre only	49	10.7

6.	Home visit	32	7.0
7.	Centre Consultative Committee	13	2.8
8.	Summer school	3	0.7

Telephone and written contact predominate as in 1 a). The lower proportions for both is probably due to the fact that many counsellors had already mentioned these means for question 1 a). Some comments on telephone contact were that they enabled direct and personal conversation, immediate response to problems, and allowed students to reveal academic worries they do not wish to raise in front of a group. The value of informal meetings in pubs is also mentioned.

### 2.3 Question 2 a)

Stated: 'The usual pattern of my study centre sessions (e.g. the rough proportion of time spent on different activities) is as follows:

To summarise, Senior Counsellors were asked to indicate how often a particular activity was mentioned by Counsellors. Thus:

		%
1.	Group discussion (counselling/tuition)	402 88.0
2.	Individual counselling/ discussion	394 86.2
3.	Using TV, radio, playback facilities	302 66.1
4.	Preparing centre/ administration	107 23.4

5.	Other activities	35	7.7
6.	No answer	4	0.9

It was pointed out that low numbers in category 4 above, are probably due to the fact that although Counsellors do this, most wouldn't think of mentioning it as part of their pattern of study centre sessions.

Patterns mentioned, though probably not wholly representative, were:

1.	Viewing TV	30 minutes
	General discussion	60-90 minutes
	Small group 'self-help' sessions and individual counselling	30-60 minutes
2.	Individual talks with early callers	30 minutes
	Group discussion	60 minutes
	Individual talks with late stayers	30 minutes

And as an indication of the pattern of activities of, perhaps, one of the less successful counsellors, a counsellor of S100 students resigning at the end of this year said:

Talking to students	0-30 mins
Writing up notes	30 minutes
Private work/meditation	90 minutes

2.4 Question 2 b)

Asked about time spent outside each study centre session recording and acting on matters raised during sessions.

Alternatives provided with results, are:

1.	$\frac{1}{2}$ hour to an hour	211	46.2
2.	More than 1 hour	193	42.2
3.	Less than $\frac{1}{2}$ an hour	43	9.4

One Senior Counsellor asked if there should not have been a question asking how much time was spent reading course units.

2.5 Question 2 c)

Stated: 'The best possible conditions for group discussion to occur at the study centre where I share responsibility for creating these discussions are as follows'.

No system of categories was devised for this question. A large number of comments were made, too numerous to summarise adequately here. They included the need for:

Comfortable informal surroundings

A core of regular attenders

A planned programme

Small groups

Stimulating broadcasts and course materials.

2.6 Question 2 d)

Stated: 'I find the best ways to encourage the less forthcoming students to contribute to group discussion are:'

No system of categories was devised for this question. There has again been a large selection of comments. They include such suggestions as:

Appealing to them individually

Introducing them to a small number of other students

Exercising firm chairmanship of discussions to prevent a few students dominating it

Asking them direct questions which they can easily answer.

It was pointed out that some counsellors mentioned mutually exclusive techniques, e.g.

Ensure that idea of participation is accepted at the outset.

Shy students should be left alone to settle in - if they feel there is no pressure to contribute, I find that they do begin to participate.

It was also suggested that lack of participation had not been a problem since those who attend need little prompting to do this.

2.7 Question 2 e)

Asked in what ways students act as independent learners at



the Study Centre. This may be felt to be a contradiction in terms, and I should imagine that some counsellors would have had difficulty in interpreting the question. Results in the categories devised were:

		%
1. Using playback facilities	162	15.5
2. In self-help groups	154	33.7
3. Watching/listening to programmes	129	28.2
4. Private study	117	25.6
5. No answer	65	14.2
6. Other ways	45	9.9
7. Not at all	29	6.4

2.8 Question 2 f)

Asked what possibilities and problems were envisaged in counselling current students in 1972.

No system of categories was devised for this question. Once again a very large number of comments have been forwarded.

Some of the envisaged problems were:

Having insufficient time to devote to 2nd level students as Foundation Course students will need most help.

2nd level students having a new counsellor may be at a disadvantage through lack of continuity.

Inability to give subject matter advice to 2nd level students with consequent disincentive to attend study centres.

Not enough students in any one course to set up self-help groups.

Problem of coping with regulations and other advice and of dealing with administration.

Some advantages for 1972 were thought to be:

Fewer problems with O.U. system after a year's experience.

Use of 2nd level students as discussion leaders for Foundation Course groups in relevant subject areas.

Likelihood of more viable groups and more social activities with larger numbers.

2.9 Question 3 a)

Asked what were the most important matters that were raised by students with the counsellor:

The results were:

		%
1.	Administrative problems	221 48.4
2.	Subject-matter problems	221 48.4
3.	Study technique	170 37.2
4.	Personal/private/domestic problems	154 33.7
5.	Assessment/marking	133 29.1
6.	Other matters	89 19.5
7.	2nd level courses	54 11.8
8.	Lack of facilities, e.g. books	15 3.3
9.	No answer	6 1.3

Administrative problems referred to matters such as error in units, late answers to queries, interpretation of regulations, delay in return of assignments and so on.

Subject-matter problems include comprehension of material, coping with several disciplines, etc.

Study technique includes essay writing, organising study time, how to revise, etc.

Personal/private/domestic problems include family illness, need for reassurance, payment of fees, etc.

Assessment/marking includes examinations, fairness of marking, summer school assessment, problems with CMAs and TMAs, etc.

2nd level courses includes choice of courses, available combinations, summer schools at 2nd level, obligations, etc.

Lack of facilities includes libraries, books, poor study centre environment, etc.

Only a small number of comments were received here. Some specific problems mentioned were:

Handling of experimental work at home

, The mysteries and miseries of credit exemptions

Work load and lack of time commonest problems raised by students

CMA grades not very helpful for learning - mistrusted even.

Other problems included exemptions, examinations, what happens at summer school, value of OU degree, Mrs Thatcher.

2.10 Question 3 b)

Stated: "The Open University provides me with certain kinds of support in helping the student. I should welcome the following types of extension to this service".

The categories devised for this question appear to have been the least adequate since the highest scoring category is "Other".

		%
1.	Other	134 29.3
2.	Improved provision of learning materials	113 24.7
3.	Improved provision of information on students	85 18.6
4.	No answer	79 17.3
5.	Improved provision of general information	71 15.5
6.	Improved contact/support other staff	71 15.5

Few comments by counsellors were received here. Among the suggestions were:

Cards saying "Sorry I haven't seen you recently. I realise you may be getting on quite well but is there anything I could do for you?"

More support for students during the summer months.

Access to H.E.K. or some of the items which are not the standard laboratory pattern.

2.11 Question 3 c)

Asked the counsellor to tick one or more of 8 groups indicated on the form with whom stronger links were felt to be required.

1.	My students' correspondence tutor/s	277	60.6
2.	My more remote students	172	37.6
3.	Other counsellors at the study centre	141	30.9
4.	Authors of courses my students are on	115	25.2
5.	Staff at Regional Office	95	20.8
6.	Authors of future courses	92	20.1
7.	Policy making bodies	66	14.4
8.	Centre Consultative Cttee.	30	6.6

Presumably the new arrangements for course tuition in 1972 will help to overcome the counsellors' desire for stronger links with the correspondence tutor.

2.12 Question 3 d)

This question asked for recommendations as to modifications to forms currently used in the counselling system. Senior

Counsellors were asked simply to record which forms their counsellors had suggested be modified.

		%
1.	C1 form	108 23.6
2.	C5 card	91 19.9
3.	C1A form	82 17.9
4.	C2 form	77 16.9
5.	C3 form	63 13.8
6.	C1L letter	58 12.7
7.	C4 form	44 9.6

A number of detailed suggestions for modification were made.

Main suggestions for C1 appear to concern the coding system.

Details of suggested modifications will be passed on to Regional Tutorial Services in a separate document.

#### 2.13 Question 4 a)

Asked counsellors what aspects of briefing and training were missing or in need of modification.

No system of categories was devised for this question. A number of comments and suggestions have been received, including:

The need for more meeting with other counsellors, tutors, senior counsellors and correspondence tutors.

More information on O.U. policy, philosophy, statistical data, e.g. on drop-outs, administrative system etc.

Advice on what to do about non-attenders, dominators of discussion groups, how much training in study-methods to give, how to use study centre equipment.

Some idea from course authors as to why courses have been arranged as they have been.

Advance access to units, assignments and information on programmes.

Plea to avoid the problem-oriented approach to counselling and to state the function of the counsellor more positively.

More detailed and specific advice on the workings of the administrative system: in particular some indication of responsibilities at Walton Hall, e.g. who is the correct Assistant Secretary to contact in a certain situation?

Employment of counsellors at Summer Schools to give first hand experience of them, and the problems students face.

Need for "inquest" on 1971 operations.

2.14 Question 4 b)

This asked for any other points on counselling a counsellor wished to bring to the attention of the Senior Counsellor or Regional Director.

No system of categories was devised for this question. Again there are a considerable number of comments. There were several comments relating to attendance such as:

Could attendance of students be an (informal) obligation rather than purely optional

Why not make it obligatory for students to contact counsellor periodically?

So long as the Study Centre remains marginal to the learning process for students, then the counsellor's academic role will continue to be nebulous. Most students would be prepared to attend an F.E. institution one night per week so why the emphasis on voluntary principle?

Only the advantages of attendance at study-centres should be stated - the optional reference was no incentive to students to attend.

Some counsellors believe it is advantageous for students to attend the study centre, at least occasionally, as discussion is as important as TMAs and CMAs. However, others do not consider this should be compulsory, although it is important that these students keep in contact with their counsellor.

Students should be firmly advised at induction to attend centres.

There is also some mention of satisfaction with the change in counsellor and tutor roles for 1972 and it is pointed out that several of the counsellors' criticisms have already been taken into account for next year, e.g. provision of course units, copy of assignment form.

A random selection of other comments will indicate the general range:

Need quicker method of dealing with students' difficulties re shortage of units etc.



Assignment form should contain a brief indication of nature of assignment for counsellor's benefit.

A small photo of individual students - as is practice in other Universities etc. - would help counsellors considerably, since contact with some students is infrequent.

Students become distressed because of lack of feedback on information on CMAs. They see their grading but do not know their mistakes.

General criticism that counsellors have been less well-informed than students.

Desirability if possible of students staying with one counsellor for several years.

Tutoring and counselling taking place on the same evening might increase attendance at study centres. Few students studying two courses can attend their study centre regularly.

Integration of First course and second level students a problem.

Feedback sent to Walton Hall by students and staff should be acknowledged.

### 3.0 Final comments

- 3.1 The data analysed for the purpose of this preliminary report is based on the summaries made of C4s by those Senior Counsellors who were able to get the information to me in time for it. Data from counsellors in certain regions could not therefore be included. A final amended version of the report will probably be made when further summaries have been received. Nevertheless the proportion of counsellors involved is sizeable, and the

addition of further data would probably not result in major changes, although a change in rank order is a possibility in some cases.

3.2 The allocation of written answers to categories is always a difficult task. The fact that Senior Counsellors may have interpreted their counsellor's answers in different ways should therefore be borne in mind.

3.3 The C4 form would probably benefit from some modification for future years. For instance, very little additional information is to be had from 1 b) over that gathered in 1 a) and 2 e) is not easily understood.

Alan Thomas I.E.T.

December 1971

ANALYSIS OF FORM C4

Counsellors' Report on Counselling and the Study Centre

Final Report January 1972

1.0 Introduction

1.1 This final report on the contents of counsellors' C4 forms supplements the preliminary report prepared early in December 1971. It may best be read in conjunction with that report. The amended statistical tables may be found in Appendix I. In the main, comments made about the tables in the preliminary report apply. Differences and similarities of comments made by counsellors and selected by Senior Counsellors is also given.

1.3 The report is based on summaries received from 30 Senior Counsellors. This represents responses from 949 counsellors (about 30% of all counsellors). Only three sets of summary sheets had not been received at the time of writing. All Regions are represented.

2.0 Results

2.1 Question 1 a)

Asked in what ways counsellors came to know their students apart from regular meetings at the study centre.

The final figures do not differ markedly from those given in the preliminary report. However, 'other written contact' now replaces 'CIL introductory letter' as the second most mentioned method. See table 2.1.

2.2 Question 1 b)

Asked what had been the counsellors most valuable means of contact apart from regular meetings at the study centre.

Final figures indicate no change from the situation indicated in the preliminary report. See table 2.2

2.3 Question 2 a)

Asked about patterns of activity at counselling sessions.

In order to summarise Senior Counsellors were simply asked to indicate how often a particular activity was mentioned by counsellors. Final figures confirm the position given in the preliminary report. See table 2.3.

2.4 Question 2 b)

Asked about time spent outside each study centre session recording and acting on matters raised during sessions.

Final figures confirm the position given in the preliminary report. See table 2.4.

2.5 Question 2 c)

Asked about the best possible conditions for group discussion to occur at a counsellor's study centre.

Factors most frequently mentioned were:

- i) Study centre environment - the quality of the study centre environment was felt to be of great importance in achieving, or inhibiting, successful group discussion. Generally favoured was the smaller type of study room, preferably warm, light, and informally and comfortably furnished, pleasantly decorated and devoid of a classroom atmosphere. The availability of refreshments was felt to be important and freedom to smoke. The study centre should be informal, relaxed and welcoming.
- ii) Attendance - a small group is preferred. Figures mentioned are not less than 3 nor more than 12. A nucleus of regular attenders is seen as useful.

Also students who are all have a common problem, come prepared to discuss some topic or problem and know each other outside study centre sessions.

- iii) Organisation of sessions - a prearranged programme which students know about in advance is often mentioned. A definite topic motivates students to attend. Also the occurrence of counselling sessions on the same evening as a TV or radio broadcast or a tutorial.
- iv) Content of discussions - discussions may benefit from a controversial topic or from topics drawn from the units. Specific problems raised by students are useful. One suggestion was the presence of a visitor such as an MP, foreign students or people working in the discipline.
- v) Role of counsellor - some counsellors think the counsellor should be an initiator and leader of discussions. Others think he should play only a background role. He could, of course, be both.

Question 2 d)

Asked about the best ways of encouraging less forthcoming students to contribute to group discussion.

A commonly mentioned method was to approach these students directly. It was suggested that the counsellor should ask tactful, simple questions, ask directly for their views and opinions, ask them to agree or disagree with some point raised without asking for a supportive statement, and to ask them for their assistance and advice.

Other points mentioned were:

- i) Keeping groups small, perhaps by dividing up into groups of about 4 where necessary. One idea to construct small groups consisting entirely of the

less forthcoming students. Other people felt that allowing the more voluble members of a group to talk themselves into an illogically usually encouraged the more timid students to contribute.

- ii) Establishing a friendly, informal atmosphere. The use of Christian names. Leaving students alone together at the start of a session to enable acquaintances to be made. Making each student feel of value.
- iii) Ensuring individual consultation with the counsellor to get to know the students' needs and interests. Asking the student in advance to prepare something. Asking him to come to sessions early.
- iv) Encouraging contact with other students. Getting another student to talk with the reticent student before the start of the session.

There were some remarks to the effect that students did not need encouragement to contribute and that less forthcoming students do not attend study centres. There was also some feeling that students who do not contribute to group discussion nevertheless benefit from attendance and that they should not be pressured into participation.

## 2.7 Question 2 e)

Asked in what way students act as independent learners at the study centre.

Final figures remain broadly as in the preliminary report.

Changes in rank order are 'in self-help groups' from 2 to 1, and 'private study' from 4 to 3. See table 2.7.

## 2.8 Question 2 f)

Asked what possibilities and problems were envisaged in counselling current students in 1972.

The main problems mentioned about counselling 2nd level students were as follows:

- i) Lack of time on the part of both counsellors and students for an adequate relationship. Also some fears of inadequate accommodation.
- ii) The general demand by students for subject-matter advice which counsellors may be unable to give at 2nd level. Difficulty of giving advice on 2nd level courses without the course materials, academic acquaintance with a large number of disciplines, and with limited time due to commitments with Foundation level students. Feeling that 2nd level students will consult counsellors largely for non-academic problems, and probably by means of telephone and letter rather than by attendance at the study centre. Consequent suggestions of the need for a 2nd level counsellor, with subject based groups.
- iii) Change of counsellor from Foundation to 2nd level resulting in discontinuity of contact.
- iv) Increased administrative, personal and academic problems among 2nd level students.
- v) Problem of persuading non-attenders at study centres at Foundation level to attend at 2nd level.

Among the possibilities that were mentioned were:

- i) The possibility of developing mixed discussion groups in which 2nd level students would help Foundation level students.
- ii) Fewer or no more administrative problems.

- iii) More self-help groups with larger numbers of students.
- iv) Personal relationships with and among students already established.
- v) Development of the telephone as an effective counselling/discussion medium.

2.9 Question 3 a)

Asked what were the most important matters that were raised by students with the counsellor.

Final figures indicate little change from the position given in the preliminary report. See table 2.9.

2.10 Question 3 b)

Stated: "The Open University provides me with certain kinds of support in helping the student. I should welcome the following types of extension to this service".

The final figures are given in Table 2.10. 'Improved contact/support other staff' has moved from 6 to 3. 'Improved provision of information on students' has moved from 3 to 5.

2.11 Question 3 c)

Asked the counsellor to tick one or more of eight groups with whom stronger links were felt to be required.

Final figures confirm the position given in the preliminary report. See table 2.11.

2.12 Question 3 d)

Asked for recommendations as to modifications to forms used in the 'counselling system. To summarise, senior counsellors were asked to record how often a form was mentioned as being in need of modification. See table 2.12. C2 is now ranked 3 and CIA 4.



2.13 Question 4 a)

Asked counsellors what aspects of briefing and training were missing or in need of modification.

With reference to modes of briefing and training, the need for more meetings was often mentioned. Participants mentioned were other counsellors at the centre, in the area, region or nationwide, class tutors, Staff Tutors, Senior Counsellors, course team members and in particular correspondence tutors. Visits to Walton Hall and other study centres were suggested, participation in summer schools, and the use of regional seminars.

Some suggested briefing and training items were teaching manuals, case-studies, simulations, training films and specimen completed forms.

Information and advice was asked for on OU policy and philosophy, precise role of counsellor, characteristics of the adult student, recommended sequences and combinations of higher courses, standards required of students, on the administrative system at Regional Office and Walton Hall, and on successful counselling practices. Also statistical data on dropouts, general feedback on feedback, information on the best study centres and their operations and from students on the use of centre facilities, on precedents, and on course materials.

Some suggested subjects for training were:

- techniques of group teaching and group dynamics;
- interviewing methods;
- stimulation of self-help groups;
- general uses of study-centre equipment
- especially computing facilities.

It was also suggested that among all the written material sent to counsellors there should be some indication of the most urgent items to enable speedy identification and implementation.

2.14 Question 4 b)

Asked for any other points on counselling a counsellor wished to bring to the attention of the Senior Counsellor or Regional Director.

The remarks made are of a very disparate nature, often reiterating points already made earlier. As one Senior Counsellor put it this question "brought forth an abundance of bouquets and raspberries, predominantly the latter"! One area frequently mentioned was that of administration. It was requested that counsellor receive all important information before the students. But there were also worried about the volume of information counsellors receive. To mitigate this a suggestion was to:

- i) indicate information requiring the urgent attention of the counsellor and that which he needs to pass on directly to students;
- ii) provide a short summary of the main points contained in any document for counsellors.

There were also some requests for faster and fuller provision of TMA and CMA results.

3.0 Acknowledgements

3.1 The data analysed and collated in the preparation of this final report was provided by Senior Counsellors, who made summaries of the contents of their counsellor's C4 forms. Their very valuable efforts in preparing quantitative summaries and selections of comments is acknowledged here.

Alan Thomas

Institute of Educational  
Technology

January 1972

APPENDIX I

REVISED TABLES

Ways of coming to know the students apart from regular meetings at the study centre			
			%
1	Telephone	780	82.2
2	Other written contact	540	57.9
3	CIL Introductory letter	513	54.1
4	Induction meeting prior to Jan 71	512	54.0
5	Home visit	207	21.8

Table 2.1

The most valuable means of contact with students apart from regular meetings at the study centre			
1	Telephone	575	60.6
2	Written contact	268	28.2
3	Other group meetings	143	15.1
4	Other means	120	12.6
5	No answer/meetings at study centre only	96	10.1
6	Home visit	69	7.3
7	Centre Consultative Committee	28	3.0
8	Summer School	5	0.5

Table 2.2

Activities mentioned as occurring at study centre sessions

		%
1	Group discussion counselling/ tuitional	840 88.5
2	Individual counselling/discussion	814 85.8
3	Using TV, radio, playback facilities	563 59.3
4	Preparing centre administration	206 21.7
5	Other activities	65 6.8
6	No answer	8 0.8

Table 2.3

Time spent outside each study centre session recording  
and acting on matters raised during sessions

		%
1	$\frac{1}{2}$ an hour to an hour	449 47.3
2	More than 1 hour	394 41.5
3	Less than $\frac{1}{2}$ an hour	93 10.0

Table 2.4

Ways in which students act as independent  
learners at the study centre

		%
1	In self-help groups	353 37.2
2	Using playback facilities	351 37.0
3	Private study	246 25.9
4	Watching/listening to programmes	225 23.7
5	No answer	122 12.9
6	Other ways	93 9.8
7	Not at all	49 5.2

Table 2.7

The most important matters raised by students with the counsellor (as seen by the latter)		
		%
1 Administrative problems	492	51.9
2 Subject-matter problems	456	48.1
3 Personal/private/domestic problems	339	35.7
4 Study technique	336	35.4
5 Assessment/marking	275	30.0
6 Other matters	135	14.2
7 2nd level courses	116	12.2
8 Lack of facilities, e.g. books	33	3.5
9 No answer	11	1.2

Table 2.9

The OU provides me with certain kinds of support in helping the student. I should welcome the following types of extension to this service		
		%
1 Other	240	25.3
2 Improved provision of learning materials	220	23.2
3 Improved contact/support other staff	180	19.0
4 No answer	173	18.2
5 Improved provision of information on students	166	17.5
6 Improved provision of general information	151	15.9

Table 2.10

Groups with whom counsellors expressed a desire for stronger links			
			%
1	My students' correspondence tutor/s	387	61.9
2	My more remote students	308	38.8
3	Other counsellors at the study centre	272	28.7
4	Authors of courses my students are on	266	28.0
5	Staff at Regional Office	207	21.8
6	Authors of future courses	204	21.5
7	Policy making bodies	139	14.6
8	Centre Consultative Committee	69	7.3

Table 2.11

Forms recommended by counsellors as being in need of modification			
1	C1 form	245	25.8
2	C5 card	180	19.0
3	C2 form	167	17.6
4	C1A form	138	16.7
5	C3 form	130	13.7
6	CIL letter	125	13.2
7	C4 form	91	9.6

Table 2.12

INSTITUTE OF EDUCATIONAL TECHNOLOGY

Tuition and Counselling Research

Phase 1b, 1972

An Examination of Drop-out, Wastage and Conservation  
of Part-time Tutors and Counsellors 1971-1972

1.0 Introduction

The group under study is all part-time tutors and counsellors who were contracted to work for the OU in 1971 (other than those who only worked at summer schools) and who actually started work. The group can be divided into two sub-groups:-

those who completed their contracted period

those who did not complete their contracted period

Each of these groups can be similarly subdivided:

those who were contracted for 1972

those who were not contracted for 1972

We are therefore concerned with four main groups, which I have designated W, X, Y and Z:

W = those who completed contracts in 1971 and who have again been contracted for 1972

X = those who completed contracts in 1971 and who have not been contracted for 1972

---

N.B: For reasons of space, tables and figures accompanying this report have been omitted.

Y = those who did not complete contracts in 1971 and who have not been contracted for 1972

Z = those who did not complete contracts in 1971 and who have been contracted for 1972.

Groups Y + Z constitute tutor/counsellor drop-outs during 1971.

Groups X + Y constitute tutor/counsellor wastage between 1971 and 1972.

These groups represent the current total loss to the system.

## 2.0 The Extent of Drop-out and Wastage

Figure 1 summarises the extent of drop-out and wastage.

### 2.1 Drop-out (Groups Y + Z)

Of the 3564 tutors and counsellors who started work in 1971 only 5% failed to complete their contracts. Allowing for illness, death, pregnancy and acts of God, it seems likely that relatively few people actually gave up.

### 2.2 Wastage (Groups X + Y)

Of the 3371 tutors and counsellors completing their contracts in 1971, 31% have not been contracted for 1972.

Of the 193 tutors and counsellors who dropped out in 1971, as one would expect 84% have not been contracted for 1972. (There is a small group here (Z) who dropped out in 1971, and have been contracted for 1972. These 30 people probably were among those forced to drop out through illness, moving home, etc.).



Thus of all those tutors and counsellors who started work in 1971, 34% have not been contracted for 1972.

### 3.0 Concomitants of Drop-out and Wastage

The various figures and tables which follow summarise the available data.

#### 3.1 Sex Fig. 2, Table 2

A slightly higher proportion of women who completed contracts in 1971 did not go on to 1972.

Almost identical proportions of men and women failed to complete their contracts in 1971.

#### 3.2 Age Fig. 3, Table 3

Wastage was greatest from the youngest and oldest age-groups although these groups constitute only 9% of all those starting work in 1971.

The 46-55 and 56-65 age groups lost proportionately more people during 1971 through failure to complete contract.

#### 3.3 Regions Fig. 4, Table 4

Wastage was lowest in Scotland where 4.8% of those starting work failed to complete contracts in 1971 and were not contracted for 1972; and 26.5% completed contracts in 1971 and were not contracted for 1972.

The greatest wastage occurred in the North-West. Here 5.7% of those starting work failed to complete contracts in 1971 and were not contracted for 1972; and 36.8% completed contracts and were not contracted for 1972.

3.4 Job Combinations Fig. 5, Table 5

Wastage appears to have been greatest from the class tutor only and correspondence tutor only groups. This probably reflects the change to the course tutor role embodying both functions.

The three job group (class tutor and correspondence tutor and counsellor) seems to have suffered most from drop-out within 1971, i.e. failure to complete contract.

3.5 Job Functions Fig. 6, Table 6

Wastage was greater from the class and correspondence tuition functions and least from the counselling function.

3.6 Functions within Faculties Figs. 7 & 8

Arts Faculty lost 39% of the people employed as class tutors between 1971 and 1972. Arts Faculty lost 31% of people employed as correspondence tutors and Science 32%.

These figures are somewhat misleading since they do not show how many people each Faculty lost, and do not include figures on drop-outs within 1971. However, it does seem likely that Science and Arts have 'shed' relatively more people than Social Science and Maths.

#### 4.0 Reasons for Drop-out and Wastage

We do not yet know why this wastage has occurred. Therefore the phrase 'was not contracted in 1972' needs to be understood as 'does not appear on the 1972 file of contracted tutors and counsellors'. We do not know whether this is as a result of the University's decision or the individual's decision.

Enquiries are therefore being made to discover, in particular, whether or not those people who completed their contracts in 1971 and were not contracted for 1972, reapplied or not, and if they did, why they were not contracted.

---

This analysis has been undertaken with the assistance of Richard Stammers, (IET/DP) who arranged for the provision of the necessary data. Fiona Arthurs and Tutor Records Office have also given help.

Alan Thomas,  
April 1972.

Tuition and Counselling Research

Phase 1a) 1972

Census of 1972 Part-time Staff

1.0 Introduction

Following the survey of part-time staff conducted in 1971, it was agreed that a similar exercise would be conducted in 1972. This would provide information of a very general kind on the characteristics of part-time staff, which would otherwise not be readily accessible to the University. From 1973 this information will be collected from all new part-time staff as a part of the University's normal administrative procedures. Over the years this will enable the delineation of general trends in the composition of the University's body of part-time staff.

2.0 Method

Although a survey had been conducted in 1971, it was decided to replace this for 1972 with a census. This would enable individual Faculties or Course Teams to have access to information on the various groups of part-time staff of particular relevance to them. A short one-page questionnaire was prepared and administered to 4200 members of part-time staff in mid-June. By the end of August, returns had been received from 3775 people, representing a 90% response. 34 of these were excluded from the analysis, mainly because no name or tutor number had been given.

For persons examining the tables it should be noted that:

- 1) For the purpose of analysis persons who were only counsellors were regarded as not belonging to a Faculty. Although there is a column headed C on Faculty tables this therefore always remains blank.
- 2) The inclusion of an individual in a Faculty table was dependent on the courses recorded for that individual. If an individual was tutoring in more than one Faculty he was allocated to one only. It was expected that only a small number of people would be affected in this way.

### 3.0 Distribution

A complete set of the tables are held by David Grugeon, RTS and Alan Thomas, IET. Sets of particularly relevant tables have been circulated to Regions and Faculties as in 1971.

### 4.0 Results

The following paragraphs summarise the general features of the tables. Obviously it has not been possible to comment on every available statistic. Where there are any noticeable changes from the 1971 situation these have been indicated although comparisons are not always possible. Exhaustive comparisons between years for Faculties and Regions have not been made at this stage. Anyone wanting more detailed information should consult the original tables.

#### 4.1 National Figures

Sex: the ratio of men to women is about 4 : 1 in favour of men. This is similar to the 1971 situation.

Age: the age distribution remains almost the same as in 1971. 83% of part-time staff are below the age of 47. About 6% are 26 years old or younger. The largest group, 48%, are between the ages of 27 and 36.

There appears to be a larger concentration of course tutors than counsellors in the younger age groups. About 60% of course tutors are under 36 as opposed to 40% of counsellors.

Place of Work: the distribution across place of work for 1972 is:

University	22.5
Polytechnic	16.1
College of Education	15.3
Other full-time	8.5
Technical college	8.1
Part-time only	7.7
Secondary school	6.5
College of Further Education	6.0
Not working	5.6
Industry	3.8
	<hr/>
	100.1

The majority of tutors and counsellors thus work in either Universities, Polytechnics or Colleges of Education. This presents little change from 1971. Other full-time work includes a range of items with small numbers in each. They include people working as advisers to L.E.A.'s, in the Civil Service, Research Institutes, Services Colleges, the Church, W.E.A., Museums, Primary Schools, BBC, Special Schools, etc.

Counsellors tend to be drawn less from Universities and more from Colleges of Further Education and Technical Colleges than course tutors.

Years of Full-time Teaching in Adult or Higher Education:

19% of part-time staff have had less than one year's experience of full-time teaching in adult or higher education. 52% have less than five years' experience. Counsellors tend to be relatively more experienced than course tutors.

Most Recent Degree: 41% of part-time staff obtained their most recent degree since 1966. 16% obtained their most recent degree before 1956. Counsellors tend to have obtained their last degree at an earlier date than course tutors. These figures are similar to those for 1971.

Previous Experience of Teaching in Adult or Higher Education: Maths have the highest proportion with no previous experience of this kind (12%). Technology have a low proportion of persons with experience of correspondence tuition (2%). In no Faculty does the proportion having experience of correspondence tuition exceed 14%.

#### 4.3 Regional Figures

Sex: Region 8 has the lowest proportion of women (9.9%), closely followed by Region 9 (10.4%). Region 6 has the highest proportion of women (22%).

Age: Region 5 appears to have a relatively high proportion of persons over the age of 66 (11%). Region 11 has the highest proportion in the 27-36 age range (57%). Region 6 has the largest proportion 26 years old or less (9%).

Place of Work: Regions 7, 19, 11 and 12 have high proportions of people working in Universities ranging from 32% to 45%. Region 9 has a very high proportion of persons working in Polytechnics (42%) whereas Region 10 has a very low proportion (2%). In Region 12, a high proportion are secondary school teachers (14%), and in Regions 3 and 5 a little over 10% work in secondary schools.

Years of Full-time Teaching in Adult or Higher Education: Region 6 has the highest proportion with less than one year's experience (27%). The lowest proportion with less than one year's experience occurred in Region 9 (10%).

Most Recent Degree: Regional distributions are generally similar to the national figures.

Previous Experience of Teaching in Adult or Higher Education: Region 12 has the highest proportion with no previous experience (12%), whereas Region 9 has the lowest proportion (3%). However, Region 9 has the lowest proportion with experience of correspondence teaching (3%).

Alan Thomas,  
November, 1972.



INSTITUTE OF EDUCATIONAL TECHNOLOGY

Report on the 1973 Counsellor Characteristics Study

1.0 Background to the Study

The study which forms the basis of this report emerged from the discussions held in the latter months of 1972 on possible approaches to research into the counselling area for 1973. Among the areas considered were those of selection, briefing and evaluation of counsellors, and it was suggested that a useful line of enquiry related to these areas would be a study of counsellor characteristics. In particular, it was thought that the general approach known as the 'critical incidents' method might be appropriate.

Any rational selection system is, of course, dependent on the ability to identify characteristics of applicants who are related to performance in the job. The aim of selection is to ensure that people who possess characteristics positively related to performance are given priority over those who lack such characteristics. The problem is to identify both indicators of performance and then characteristics which are related to the indicators. Very crudely the factors entering into the equation are:

Actor's properties  $\longrightarrow$  Actor's activities  $\longrightarrow$  Outcomes of action

If it were possible to identify unambiguously the preferred outcomes, it might then be possible to identify 'properties', such as sex, or age, related to the preferred outcomes. Unfortunately, attempts to identify characteristics important to, for instance, teaching

performance, have usually proved inconclusive despite years of effort and expenditure running into hundreds of thousands of pounds. It would be prudent, then, not to expect too much from our own more modest efforts.

However, this study has tried to increase understanding of what is involved in counselling as it currently exists, by trying to identify some important ways in which counsellors' behaviour varies. The research has tried to establish some qualitative dimensions of counsellor activity which seem to be important to performance. These dimensions are ones which may be taken into account, and often have been in relation to counsellors cited by senior counsellors, when a judgement about the effectiveness of ineffectiveness of counsellors is made. The dimensions given largely refer to the activity part of the above equation, rather than background characteristics, and therefore highlight some characteristic ways of behaving which seem to differentiate counsellors who are regarded as effective from those who are regarded as ineffective.

Since these dimensions were identified via the critical incidents approach, it is necessary to review briefly the elements of this.

The critical incidents method is an approach to the empirical description of some activity which was developed by J.C. Flanagan and his associates during the last war. Since that time it has been applied in a wide range of contexts, including studies of nursing, teaching, and research work. The main sources consulted

during the design of this study of counselling were Flanagan's original paper and two more detailed accounts by Jensen.\*

In outline, Jensen's approach to the study of teaching is this. He notes that there is often disagreement about the qualities, traits and characteristics which contribute to effective teaching, and that there is a need for a better understanding of the patterns of personality traits and behaviours of teachers.

He suggests that the 'critical incidents' approach can give important information about the kinds of behaviour that count as elements of better or worse teaching. Basically the approach involves identifying and differentiating those kinds of behaviour which are commonly agreed to constitute good teaching from those which constitute poor teaching. The outcome is a list of 'qualities' which can be used as a basis for the evaluation of teaching performance and as a guide in the selection process..

In order to generate information about these 'qualities', selection of teachers, trainee-teachers and administrators were asked to provide descriptions of things they had seen teachers do which had made them decide that that person was an effective or ineffective teacher. These descriptions, called 'incidents', were

\* FLANAGAN J.C. (1954). The critical incident technique. The Psychological Bulletin Vol. 51, No. 4.

JENSEN A.C. (1951). Determining critical requirements for teachers. Journal of Experimental Education 20.

JENSEN A.C. (1960). Determining critical requirements for nurses. Nursing Research 9.

then classified in terms of the qualities of the actor reflected by them, and descriptive statements summarising each category produced. These statements then constituted a description of the components of effective and ineffective teaching. In other words, they described some important things which teachers did which were associated with being judged as more or less effective.

The conceptual basis of the critical incidents approach poses several problems. The definition of an incident, for instance, as "any observable human activity sufficiently complete in itself to permit inferences and predictions to be made about the person performing the act" is far from helpful. A seminar on the approach held in IET in June 1973 did not go far towards solving these problems, so the approach adopted represents a compromise between the general principles of the method and the special peculiarities of our particular situation.

## 2.0 The Structure of the Study

The aim was to generate descriptions of counselling behaviour which appeared to be either effective or ineffective by collecting specific examples of such behaviour. Initially it was planned to collect such examples from students, counsellors, and Senior Counsellors but limitations on resources and the time needed to work out a suitable approach to students and counsellors, meant that only Senior Counsellors were included. However, they seemed a particularly important group since:

1. They had a major responsibility for selecting and supervising counsellors
2. They would be in the best position to have comparative information about the work of a range of counsellors
3. They would probably know a good deal about what counsellors actually did.

The approach to Senior Counsellors took two forms. It was thought that the best way to obtain information would be by interviews, but it was clear that there were too many to allow this in every case. It was therefore decided that one Senior Counsellor in each Region would be interviewed and the remainder contacted by post.

The Senior Counsellors were selected for interview randomly within certain constraints. The two involved in the first trial interviews were selected on the basis of previous acquaintance and because of their expressed willingness to co-operate. Senior Counsellors who were relative newcomers to the University were excluded from the study, to ensure that reports would only come from people with some considerable experience of working with counsellors. However, information about the study was sent to these people to let them know what was going on. Finally, interviewees were selected from those who had replied to an earlier request for information about counsellors (separate report in progress) since it was assumed that these would have no objections to talking about counsellors they thought were ineffective.

Reports sheets were sent to all other Senior Counsellors.

The distribution of respondents were as follows:

	Number of Senior Counsellors		
	Interviewed	Returning Report Sheet	Total
R1 London	1	1	2
R2 South	1	1	2
R3 South-West	1	1	2
R4 West Midlands	1	2	3
R5 East Midlands	1	1	2
R6 East Anglia	1	2	3
R7 Yorkshire	1	-	1
R8 North-West	1	1	2
R9 North	1	-	1
R10 Wales	1	-	1
R11 Scotland	1	3	4
R12 N. Ireland	1	-	1
Total	12	12	24

A total of 24 out of a possible 39 Senior Counsellors gave information.

## 2.1 Interview Format

All the interviews were conducted by the author. Most were held in the Regions with a few at Walton Hall. The interviews lasted between 1 and 3 hours and all were tape-recorded with the respondents' permission. Recording was felt to be essential in view of the length of the interviews and their open format.

In each interview the respondent was asked to talk about four counsellors in turn, two of whom were thought to be outstandingly good and two regarded as outstandingly poor. The terms 'good', 'effective', 'successful', and 'poor', 'ineffective', 'unsuccessful', were used fairly loosely and were assumed to index in a general way opposite poles of a qualitative dimension. The term 'outstanding' also created some problems but was intended to refer to someone who stood out in the respondent's mind as a better or worse counsellor. It was assumed that Senior Counsellors would have made judgements in their own minds as to the relative merit of their counsellors, and the idea was to talk about some who had been judged in a generally negative or positive manner.

In Jensen's use of the critical incidents approach, respondents are asked to describe 'the final thing' which led to a judgement of effective or ineffective. This seems misleading since the final thing may not reflect the trait which underlies the judgement. Senior Counsellors were therefore asked for descriptions of things counsellors had done which illustrated their effectiveness or ineffectiveness. I also asked for a few brief background details for each counsellor, though the identities of particular individuals were kept anonymous. Therefore, instead of simply collecting isolated descriptions an attempt was made to generate a general picture of the characteristics of each individual.

## 2.2 The Report Sheets

The report sheets tried to reproduce the format of the interviews on paper. Senior Counsellors were therefore sent four sheets each, two for effective counsellors and two for ineffective counsellors. As was expected, they did not provide as rich a source of information as the interviews, and only about half of those approached replied.

## 2.3 Method of Analysis

When all the interviews had been completed, the recordings were transcribed and extraneous material (such as introductory remarks) eliminated. A summary card was then prepared for each counsellor who had been considered in the interview which listed the characteristics ascribed to the counsellor by the respondent, and those which were reflected by the illustrative examples given. This process was repeated with the report sheets. In all, summary cards were prepared for 45 effective counsellors and 30 ineffective counsellors. The lower figure for ineffective counsellors reflects the fact that some Senior Counsellors did not regard any of their counsellors as being generally ineffective. About 600 summary statements were processed.

A preliminary list of categories of effective and ineffective counselling behaviour was then produced and each statement on each summary card was coded into the set of categories. The list of categories was then revised and rechecked against the original interview transcriptions and report sheets. Finally,



a check was made to see how many Senior Counsellors had cited examples of counsellors whose behaviour could be seen as falling under each category.

### 3.0 Problems Encountered During the Study

One problem concerns the conceptual basis of the critical incidents approach. The strength of any research approach varies with the context in which it is applied, and it is probably fair to say that the critical incidents approach is not at its strongest as applied here. In an ideal situation it would be possible to identify good and poor performers of some activity by reference to some measurable output, as, for instance, with some simple manufacturing task. Then observers could be asked to observe good and poor performers in order to identify behaviour which contributed to the outputs.

In the case of counselling, one is faced with a complex interactive activity where 'outputs' cannot be identified or measured without great difficulty. The counsellors cited as examples of effective and ineffective counsellors have been selected on the basis of the varying criteria and knowledge of individual Senior Counsellors. The relevance of the characteristic activities of such counsellors is related to those individual criteria, but they are assumed to have a common core. Moreover the list which emerged encompasses a range of activities, some of which may apply in every case and none of which would occur in all cases. The characteristics represented in the list thus represent those believed to contribute to effective or ineffective counselling, based on the routine observations and experiences of Senior Counsellors.

Secondly, the specific examples of counsellors' behaviour given by Senior Counsellors were not always highly detailed. Sometimes respondents found it difficult or impossible to give particular examples of the sorts of things individual counsellors had done. Therefore general statements about the activities of particular counsellors have been included in the analysis, and in fact more detailed examples were often illustrations of the general statements.

Thirdly, the precise formulation of the final categories posed some problems. These included how many descriptive statements to put together as a single general category, and the problem of overlap between categories. These problems have been dealt with in a fairly intuitive manner.

#### 4.0 Characteristics of Effective and Ineffective Counsellors

The following list gives the characteristic activities contributing to effective or ineffective counselling derived from

the above analysis and based on information provided by Senior Counsellors. The order is arbitrary.

Characteristic Counselling Activities

EFFECTIVE

1. Shows enjoyment and enthusiasm for the job. Has a positive approach and regards the job as important.
2. Shows initiative and independence in action. Is self-sufficient and accepts responsibility for helping students.
3. Shows sympathy and empathy in relations with students. Is able to see situations from the student's point of view.
4. Shows liking for and interest in students, and encourages contact.
5. Has an outgoing manner and is relaxed, informal and confident.
6. Keeps informed on the circumstances and progress of all students, and is able to respond positively to Senior Counsellor's requests for information.
7. Identifies and pursues individual students' problems, giving practical constructive advice.
8. Refers problems to Senior Counsellor when necessary, providing clear and detailed information. Considers and acts on advice given.
9. Devises flexible programme of Study Centre activities. Prepares for sessions in advance.
10. Creates and maintains group activities. Manages groups constructively and with tact, encouraging students' contributions. Creates relaxed informal atmosphere.

NEGATIVE

1. Is unenthusiastic about the job and does not derive enjoyment from it. Regards the job as a chore.
2. Is passive and lethargic. Fails to accept responsibility for action.
3. Lacks sympathy for students. Is unable to appreciate student's situation.
4. Is remote, detached and uninterested in students and does not encourage contact.
5. Is withdrawn, reserved, formal and lacks confidence.
6. Does not keep informed on the progress and circumstances of students. Cannot give Senior Counsellor information when asked.
7. Does not identify or pursue individual students' problems. Ignores problems or fails to give constructive advice.
8. Passes all problems to Senior Counsellor or fails to consult Senior Counsellor at all. Fails to provide clear, detailed information. Does not respond to advice.
9. Lacks programme of Study Centre activities. Treats sessions as disconnected parts. Works 'off-the-cuff' without preparation.
10. Unable to initiate or maintain group activity. Dominates student and has patronising approach. Does not encourage informal discussion.

- |                                                                                                                                                                                                                                                                                                                                                                                                      |                                                                                                                                                                                                                                                                                                                                     |
|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| <p>11. Shows knowledge of subject-matter(s) and adapts approach to OU context.</p> <p>12. Knows administrative regulations and brings them to the attention of students when necessary.</p> <p>13. Co-operates with part-time colleagues. Seeks to co-operate and liaise with tutor(s), relating own 'tutorial' work to tutors where possible.</p> <p>14. Attends to management of Study Centre.</p> | <p>11. Shows ignorance of subject-matter(s). Does not adapt approach to OU context.</p> <p>12. Is unaware of administrative regulations and cannot advise students accurately.</p> <p>13. Ignores part-time colleagues and fails to liaise with tutor(s).</p> <p>14. Does not execute responsibility for managing Study Centre.</p> |
|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|

In order to expand and illustrate the meaning of these categories, examples of the kinds of activities reported by Senior Counsellors follow.

#### Category 1

A characteristic frequently mentioned in relation to effective counsellors was 'enthusiasm'. Counsellors were reported as being 'emotionally and intellectually committed to the Open University', 'interested and enthusiastic', 'deeply interested in the OU system' and 'enthusiastic about the philosophy of the University'. As one Senior Counsellor put it:

"I can rely on this counsellor never to let a student fall by the wayside through neglect. This is a most important characteristic. Other qualities can be improved on, but not this one. We can't replace these people. They came in because they believed in the OU and it's their conscientiousness which keeps them going."

Another said:

"This counsellor is concerned with adult students and their problems, and with the philosophy of the University and doesn't count the pence."

Such enthusiasm and interest was indexed in numerous ways, such as the amount of effort given to the job, willingness to participate in socials, and Study Centre and Regional Consultative committees, positive contributions to briefing sessions and even authorship of journal articles about the OU.

Ineffective counsellors were sometimes reported as being 'uncommitted', 'lacking in interest' and 'unable to derive satisfaction from the job'. Remarks made by Senior Counsellors included:

"He doesn't take the job seriously. It is a routine for him and he treats it like an administrative chore."

"I don't think he enjoyed it. I was immensely relieved when he didn't re-apply. I could see absolutely no case for keeping him."

"He wanted nothing from the OU but money. He was obviously uninterested in the job and didn't look for satisfactions."

"Lacks commitment and conscientiousness. He gives the impression that his heart is not really in his OU work. While he, more or less, 'goes through the motions', he certainly would not do more than is necessary to keep up appearances."

"There appeared to be a tendency to regard the post as a sinecure to be maintained by good relations with OU senior staff."

"He is always full of educational theories, but they are really intended to explain why counselling can't work - he has no theories on how it might work better."

## Category 2

Effective counsellors were often reported as 'taking the initiative' in their work and as being 'independent' and 'self-sufficient'. Often examples were given of counsellors going 'beyond the call of duty' in their efforts to help students, and this included:

- going beyond time in Study Centre sessions;
- giving extra sessions without pay;
- encouraging students to meet at the counsellor's home;
- visiting students at home, without pay.

Further examples of behaviour within this category are:

"He fits in 3 or 4 special sessions a year for which he asks no payment. They are short Saturday sessions held at his own University's laboratories, or places like this, where he can produce enrichment activities for the biological sections of S100. For instance, he demonstrates instruments that students wouldn't normally see. And he gets a good attendance for it."

"The counsellor went on his own initiative, without payment, to help a disabled student with some material. This indicates to me that he is keen for his students, is not particularly concerned to make money out of the system, and had the initiative to do it on his own".

"He will go to endless trouble to provide things for students. He managed to get a film of Hamlet from London and then booked a theatre to show it in."

"He is very self-sufficient but contacts me when necessary. He will phone Walton Hall direct if he thinks it is the best way".

"When his students get into trouble, he gives a lot of his own time and takes it very seriously. One woman is still in the University only because of his efforts in his spare time trying to help her sort out a difficult family problem."

"He does revision sessions even through the summer school period. He decided that the best way to get things out of S100 students was to see what their own expertise was. He found that many of his students had scientific occupations, such as hospital and laboratory work. He organised a series of one-day trips where they all visited the work-place of a student so that they could see the application of some piece of scientific work. People like talking about their work so it enabled students to build up their confidence when showing the others round. He did ask for payment, but accepted it when we couldn't."

"He issued monthly programmes. In these, I was astonished to see things like Saturday visits to computers and power stations. He got no pay for this. He organised two socials alone, and tried to set up a discussion group. He showed the kind of concern, initiative and involvement which I think is a part of good counselling."

Ineffective counsellors, by comparison, were reported as being passive and lethargic. As one Senior Counsellor said:

"He has learnt that he can lie low, say nothing, do nothing and simply draw the money. He is totally negative and makes no attempt to go out to students. He waits for them to come to him, and on some occasions discourages students from asking him questions and bringing problems to him. He regards the job now as a sinecure."

Another reported this example:

"At the time when students were upset at CMA-handling problems, he sent me a note on a C2 asking me to reassure 3 students that their grades would eventually be recorded. This at least characterises the man's lack of drive and initiative."

Similarly:

"He is unable to act, to think of constructive solutions to problems which he could analyse and describe with great acumen."

"He lacks flair, energy and initiative".



Category 3 and Category 4

These categories have been taken together, since although it may be convenient to list them as separate items, the examples which illustrate them often seem to embody both qualities.

A frequently mentioned characteristic of effect counsellors is 'sympathy'. This was often referred to as having 'concern for students'. To describe precisely how this sympathy was manifested was not easy. Thus:

"It was the way in which students approached him and he them. Good counselling rests on mutual trust, liking and friendship between student and counsellor. You can tell if this relationship is there when a student meets his counsellor. It is there in his case. They were glad to see him and he them. It may have been something like greeting a student by his christian name, the type of smile he gives them, the type of gesture and manner in which he greets them. It's all very difficult to describe but it's simply a pattern which indicates warmth and friendliness".

"In terms of personal, face-to-face contact he can immediately establish a warm relationship which does a lot for students' confidence. I can't specify it, but you do learn when reappointing people that this is the key factor in the counselling role, the ability to establish and maintain relationships with adults fairly quickly."

Other examples were:

"He has a sympathetic attitude which enables students not to pour out their hearts but to discuss the problems".

"He obviously has the interests of his students very much at heart. He enjoys his contact with students and forms warm relationships with them."

Other aspects of this characteristic were revealed in the way counsellors dealt with individual students' problems and in their management of groups, both of which will be discussed below.

Ineffective counsellors sometimes appeared to lack sympathy for students and did not encourage contact. Again this was reflected in their handling of problems and group work. One Senior Counsellor reported:

"I visited the study centre and she approached me and said, 'Ah, here's Mr. X from the Regional Office. How pleased we are to see you. You must come and talk to Mrs. Y'. Mrs. Y is about 42, married with a family, and the counsellor was talking to this woman as if she was a schoolgirl, much to her embarrassment and mine. 'Now just tell your problems to Mr. X. He will help you I'm sure.' I thought she must go."

The counsellor must have the confidence of the student and not embarrass him."

Another reported:

"His manner was slightly patronising. It was like a primary school. There was a fair amount of heavy paternalism. There was a revolt. The students wanted to discuss CMAs and he didn't. They had to appeal to me as an impartial referee."

The same Senior Counsellor said of another counsellor:

"He never extended his sympathies to put himself in the position of the students, involved in the OU machine. He ticked the register at the start and then went through the week's unit in a strong, didactic way. He had only a crude appreciation of the problems students might have in the OU. He said, 'They only come in for the tutorial. Apart from that they don't have any problems'. Although he was able on the course, he didn't have care and interest in students."

Another Senior Counsellor reported:

"Two students turned up at the Study Centre at half-hour intervals, obviously not expecting to find anything going on. They were simply looking in to see the counsellor. I tried to discuss the situation with the counsellor, and he said, 'They're all getting along all right. No-one has any problems.' He had the problem approach - 'as long as nothing shows a signal I don't need to do anything.' "

#### Category 5

Effective counsellors were sometimes described in the following terms:

confident in own ability, relaxed manner, outgoing.

Ineffective counsellors were sometimes described as having a tense, overformal approach, being timorous and lacking in confidence, reserved or withdrawn.

Category 6

Keeping in contact with all students was often mentioned in relation to effective counsellors. One Senior Counsellor put it as follows:

"She doesn't regard the study centre session as a kind of extra-mural lecture she must give each week and then tediously phone up those who don't attend. She sees the study centre session as a shorthand way of seeing 20 of her 30 students. But she also contacts the rest. If I ask her about a student, she can give several months detail of every individual student, even ones she's never seen or who have never responded to her attempts to contact.

She demonstrates that she knows a lot about her students, and I know with absolute confidence that if I phone her about a student she will know all about him. I expect every counsellor to be able to say, 'Mr. X - five letters written, no replies. Wrote to you on a C2 etc.' The counsellor must know about the student even if he only knows he knows nothing about him."

In contrast, an ineffective counsellor was described as follows:

"When I asked him how his students were getting on, he could hardly identify half of them. Good counsellors manage to build up a mental picture of all their students, even when they do not see them often. They will run through

their list, saying 'John Smith', (they know the christian name) 'yes, I've only seen him once - he came in to the induction meeting - he's a tall, television engineer - he gets on very well. Bs and Cs mostly - I believe he goes to the self-help group at Peter Jones' house - no problems there', and so on, with confidence and familiarity."

### Category 7

The ability to give help to individual students with particular problems is, of course, partly dependent on the ability of the counsellor to keep in contact with them. Effective counsellors, however, were able to pursue problems when they arose and act constructively to solve them. Here are some examples:

"She has taken up many problems with individual students. She has two remote students, both with problems. One had lost her husband and didn't know whether to carry on. The other is insecure and neurotic and has problems with essay writing. This counsellor has pursued these problems and kept in close contact with me, the Staff Tutor and the tutor. She is concerned for the individual learner".

"The counsellor visited a student who has multiple sclerosis. He had such a good relationship with his students and the Consultative Committee that he could explain the problem to them, and say, 'Look, can we organise something. This chap needs people.' Through the strength of his personality, people are now going out to visit this student."

"He doesn't try to flannel students. He doesn't offer half-baked psychoanalytical explanations of people's behaviour. He throws problems back to them, having helped them to analyse it. He regards students' problems which interfere with learning as something to be disposed of, not consolidated, wrapped in cotton wool and presented."

"Students go to his sessions in large numbers. I get quite a lot of paper from him on involved personal points offered by students. He gives people who turn to him for advice enormous confidence that they are in capable hands. He doesn't just offer sympathy and intuitions. He gives coherent, rational advice.

For instance, if a student says he has marital or housing problems, sympathy isn't enough. This counsellor will give the address of the Marriage Guidance Council or the local housing manager. He then looks at the student's record and writes them a work programme for the next month. Few counsellors do this. It's tremendously useful for the student who is, say, in a muddle with his TMAs."

"Out of a group of 40 A100 students, she detected 4 who needed some remedial work on essay-writing. She contacted Regional Office and we arranged a special tutorial. It took energy and persistence to detect these individual needs. With a group of 40 it would have been easy to stop at the group approach, and forget about individuals and their particular needs."

"The counsellor has a student who is severely disabled. He has studied the case in depth and has formed a very good, professional relationship with the student. The student - as tends to be the case with many disabled students - is not always easy to help, but the counsellor has persisted and has alerted the Regional Office to special difficulties. A special tutorial provision has been made, and special arrangements made for summer school. The counsellor has made contact with local agencies when necessary, has visited the student's home, and has enlisted the support of the student's wife.

This close relationship and detailed knowledge of the student's particular difficulties seems absolutely essential if the student is to be helped effectively. The student knows he has someone who is fully aware of his situation, to whom he can turn for guidance and support".

"A student with a spastic child, and unable to attend the study centre, was falling badly behind and contacted the counsellor about it. The counsellor visited her and decided she required some extra tuition to enable her to make up the lost ground. Some of the tuition she gave informally herself. She also contacted me and arranged for a special tutorial for this and another student who had been sick."

"The counsellor met a potential student, a recent immigrant employed in the area and living in digs. He gave good advice as to the advantages of OU study, on aspects of course choice, procedures for application and as to possible preparatory study.

He maintained contact with the applicant and helped him with preparatory work. When the student began the course, the counsellor helped him to settle in as a member of the study centre group and arranged for the use of laboratory space so that all the students could use home experimental kits together under very favourable conditions."

Partly the constructive attack on such problems involved knowing about and using the resources available: services of the Regional Office and its staff, the tutors, students, consultative committees, and Walton Hall.

Ineffective counsellors often did not know about the problems of individual students or did not do much to give help. Here are some examples:

"He hasn't done much more than tabulate problems and it isn't easy to get him to follow things up. I can remember at least two occasions when I had to write to him asking him to look more deeply into things. He was just getting messages from students, putting them on paper and sending them to me. I had to push these back and ask him to find out more. It seemed to him that this was what the system required - Walton and the Regions would solve these problems whilst he got on with something else. I think this was the sort of thing that made me lose confidence in him as a counsellor."



"She did have compassion. She visited a disabled student. But she was the kindly, compassionate, mothering-smothering kind. She did nothing practical. It wasn't good enough. She wasn't tuned in to the adult situation."

"I went to the study centre and the counsellor showed me two letters from students, indicating the need for urgent action by the counsellor. A third student was in difficulty and the counsellor had the background information. But nothing had been done."

"A student approached the counsellor in March requesting summer school excusal. The first I heard of it was when the student phoned the Regional Office to ask if excusal had been granted. I asked the counsellor why he hadn't forwarded the application and he said there had been a misunderstanding. He said the student only wanted an authorisation for late submission of assignments and not summer school excusal. I asked him to contact the student to make quite sure of this. Several weeks later the student again phoned to ask if summer school excusal had been granted. I contacted the counsellor immediately, explaining in detail the appropriate procedure, and asked him to investigate and report. I emphasised how urgent it was in view of the summer school fee liability. It wasn't until June that the application arrived at the office. The counsellor seemed unrepentant and found it hard to see what all the fuss was about".

"The counsellor had a student handicapped with arthritis and with other difficulties, but who was a regular attendant at the study centre discussion group. A fortnight before the examinations and just before the counsellor went on leave, he sent a vague C2 indicating that she would take the exam with the others, at the examination centre, taking it for granted that extra time could be arranged there. When I contacted him about this he asked for the student to be able to take the exam in her own home. This turned out to be a half-finished bungalow where her three young children had to be considered. In his absence, last-minute arrangements had to be made for a room and a colleague of his to invigilate.

Most ineffective counselling has this aspect of a lack of concern for the individual difficulties of a student".

#### Category 8

In the case of effective counsellors, the referral of particular problems to the Senior Counsellors was often mentioned. These referrals needed to be well documented and be non-trivial. Ineffective counsellors sometimes either passed on everything to the Senior Counsellor, or never referred anything. In the latter case, it was sometimes suspected that the counsellor was not identifying problems which should have been identified. When the Senior Counsellor gave advice, an effective counsellor would consider it and act on it. An ineffective counsellor might ignore any response from the Senior Counsellor.

One Senior Counsellor put it like this:

"One can judge quality by the way in which a counsellor sends in C2s. There are some counsellors who send in shoals of C2s for all sorts of reasons, many of which are not very good. Others send in none when one is fairly sure they should have done. Good counsellors send in working C2s, ones which have some point. They will occasionally send in a report saying, 'Things are going well. I've lost two students so far. The rest have been getting such-and-such grades, making use of the Study Centre equipment .....'. They do this off their own bat without having these things dragged out of them. It's a matter of using C2s appropriately."

Here are some other examples:

"She has a diagnostic approach. When she sends me a C2 she analyses the problem as best she can so that when I meet the student I feel as if I know him already. Poorer counsellors never think about contacting the Senior Counsellor. I am responsible for these students so I like a counsellor who keeps in touch with me and who makes suggestions."

"He is above average in terms of his self-briefing. I don't get many C2s from him, but when I do it's a legitimate, genuine one. For instance, with summer school excusal cases he will document them properly and give me all the necessary information. This enables the rest of the machinery to grasp the situation easily. Poorer counsellors send in vaguely worded C2s. You then have to go back to the counsellor and find out more."

"A student had a back injury which meant he was housebound for a month or two. The counsellor quickly assessed all the complications arising from this and alerted the Senior Counsellor, who made a lengthy home visit. The counsellor had judged that the case was too complex for a part-time member of staff to solve at a distance and made a correct referral."

"This counsellor has been prepared to contact the Senior Counsellor about matters over which he really needed assistance, advice or a second opinion. He has very rarely contacted me over what I'd consider an insignificant matter."

Some examples referring to ineffective counsellors:

"He hadn't read the B.A. Degree Handbook, so I had to send him replies to C2s saying, 'See BADH page 53'. Usually one such cutting reply is enough. A counsellor usually won't send you another C2 like that. But he sent innumerable ones."

"He has rarely come to me with students' individual problems. There was an instance where I knew from other sources that a student was in trouble but he didn't indicate this to me. I know one of his students who is elderly, has health problems, limited previous educational experience and difficulty in coping with the work. This only came to my notice when the student changed study centres, not before."

"He failed to obtain additional information from a student applying to the Student Hardship Fund. As time was running out I eventually contacted the

the student myself. A better counsellor would have visited or phoned the student and provided supplementary information on a C2. He would have endeavoured to collect full information and have pestered me until he was satisfied that the case had been satisfactorily resolved."

"He always phones the Senior Counsellor when he has a problem, even when the answer is in the B.A. Degree Handbook or some other publication."

"He refers all problems (which are squarely in a counsellor's province) to the Senior Counsellor almost as a reflex."

#### Category 9

Having a flexible programme of study centre activities was often mentioned for effective counsellors. Sometimes the Senior Counsellor insisted that all their counsellors produce such a programme. Preparation in advance for study centre sessions was also mentioned. Thus:

"She aims at advanced planning and devises a structured activity for each block of the course. She has set up project groups and always prepares well."

"He programmes and structures counselling sessions very well. He gave the Senior Counsellor a copy of his programme and distributed it to the students at the first session. The programme outlined possible activities for every night of the year and indicated units and assignments for discussion."

"He is the only counsellor I know who plans his work really thoroughly and in detail ahead. But he is flexible and will alter it if circumstances change."

"At the beginning of the year, in consultation with the tutor, he planned out a provisional programme of counselling and tutorial activities for the year. This was referred to the Staff Tutor and Senior Counsellor for approval. A typewritten 'agenda' was prepared for every session, and - allowing for flexibility when necessary - he made sure all the topics were covered. This 'agenda' is prepared in consultation with the students. There is no undue rigidity and his aim is to ensure that the wishes of students are met, and the limited time available is used as profitably as possible. This definite framework is welcomed by students: they know what to expect on any given evening and are able to prepare accordingly.

A worse counsellor would have adopted a completely laissez-faire approach. It would have been left to chance to determine which topics/activities would arise; students would not know what to expect and would not be able to prepare. No attempt would be made to foresee likely areas of difficulty at different stages of the course or year."

"In January he arranged his students into 'course groups - A100, 2nd Level Arts, Educational Studies and 2nd Level Social Science courses. He then drew up a broad study centre programme - A100 students every other Wednesday, Educational Studies one Wednesday per month and so on. This he sent to all his students as a suggestion. This was well before the first week of the study centre opening and in advance

of my suggestions on what should be offered to 2nd Level students."

With poorer counsellors, specific mention of lack of a programme or preparation was mentioned less frequently but was nevertheless evident from the other information given.

Two examples may serve as illustrations.

"He didn't plan his work ahead with his students. He would wait until three or four arrived and then pick up a unit and generate a bit of discussion, assuming he knew a bit more about the subject than they did. This is what the worst counsellors do. Many counsellors fall back on this occasionally, but he did it frequently."

"I went to the study centre and was introduced informally to the students. About eight or nine students were present, and the counsellor began by asking if anyone had any problems. I thought this was an ineffectual way to start. Nothing was being followed up from previous weeks, and there was no programme of activity prepared by him or them. One or two personal problems emerged and then it just degenerated into a lot of small groups talking amongst themselves. There was no structure at all to the activity. The time was rather wastefully spent for two-thirds of the people present."

#### Category 10

The ability to initiate and manage group activities also appeared to be an important attribute distinguishing effective from ineffective counsellors. Many illustrative examples of counsellors

in group situations were given. This does, of course, partly reflect the fact that Senior Counsellors were most likely to observe counsellors at the study centre, where group-work is a visible activity.

Some examples of effective counsellors follow.

"When I saw him, there were S100 and post-Foundation Level Science students in twos and threes. He was moving round with a word here and a joke there, asking how they were getting on and what they were doing. It was a fairly subtle, non-directive teaching job, helping people clarify their position and forcing them to think through it. He was well up with S100, so he had no problems there. He could bring a perspective to bear that the students didn't have at that time.

I don't think he had deliberately organised these groups. There was a very informal atmosphere. It was always first-name terms with no question of status or style of address, and always a pint in the pub afterwards. Probably he didn't always run his sessions like this, but he seemed to be capitalising on a situation he found in the room."

"In the music section of A100 he discovered he had a musician in the group. The counsellor wasn't happy with his own musical knowledge, so he handed the group over to the student. The student brought in instruments and records and talked about them. At one point the student was talking about the sonata form and the counsellor gently pointed out that the sonnet has some identical features.



The counsellor had been able to assess that the student could give the talk and had built up his confidence to help him. The he had left things alone until he could make the one contribution which everyone could learn from and that he alone could make."

"I've seen him with a group of students operating at the study centre. He has a very informal approach. They'll all be chatting around the coffee table. I felt no tenseness or uneasiness and I feel that's a nice situation. He doesn't dominate his students. He treats them as adults. He has a freewheeling attitude and lets them talk about anything. But he is prepared and knows his material backwards."

"This counsellor had A100 and A202 students. This evening there was to be an A100 tutorial with the A202 students coming in for their regular session. The evening had been planned in advance as always and the students knew what the arrangements would be. The A100 group would watch the TV programme, have the counselling session and then go to the tutorial. At this point the A202 session would begin.

The students discussed ways of tackling the next assignment. The counsellor was the chairman and he asked questions, clarifying the students' ideas and setting them into new trains of thought. He made sure that everyone contributed and livened up the session with anecdotes.

The A202 session was on a unit which the counsellor already knew something about. He talked a lot but I thought the students were getting a lot out of it. He always pointed out when he was giving his own personal views and left it open for students to form their own."

"The counsellor arranged for his Foundation Course group to come in early, so that he would be able to discuss work with them before the arrival of a higher-level self-help group, which was having its initial meeting. A third group (not one of his) arrived at the same time. He found a room for them, got his own self-help group running, clarified some points for the Foundation Course group and had all three groups settled satisfactorily with a minimum of fuss. He did this fairly quickly and without apparent exertion of 'authority'."

"A Foundation Course group of twenty-five students was divided into two sub-groups and the counsellor moved between the groups. One group always discussed the current unit. The other group discussed wider topics or perhaps a future TMA. Students were free to move from group to group each week.

The aim was to try to create something like a self-help group, but which was very much under her control. She guided the groups and made sure they were not wasting time. As the groups became more confident and effective, the counsellor withdrew and made herself available for consultation by individual students."

"He has formed an extremely effective T100 counselling group. Students from engineering and social science backgrounds, men and women, were able to hold effective discussions relating to the technology course. He is a good discussion leader. He also started a self-help group which has continued throughout the year."

"He watched a T100 TV programme with his group of Foundation year students. Immediately after the broadcast there was silence. Then a student asked a question about an aspect of the programme. This appeared to offer the counsellor the opportunity he wanted. He immediately, though in a kindly way, turned the question back to the group to encourage them to take some steps towards answering it. He was concerned to get the group to use its own resources, to avoid any tendency to lean on him in the subject-based discussion and in general to step further down the road to independent learning.

A worse counsellor might have answered the question without perceiving the opportunity which it afforded. This might have produced a prolonged exchange between that student and the counsellor, to the detriment of group participation. More seriously, a chance to set the group to using its own resources on the question would have been missed."

"An Italian lady was a member of his group. On this occasion, she said nothing during the discussion and as time passed I began to be concerned about this. Just as I thought nothing would be done, the counsellor, in a perfectly natural way, asked, "What do you think ...? (her first name). She then took part in the remainder of the discussion. The counsellor explained later that she was very reserved, and that he had been careful not to attempt to get her to contribute too early for fear she might be embarrassed."

"On one of my visits to the study centre, I was pleased to note that, apart from a brief glance of recognition, he made no attempt to introduce me or disturb the flow of discussion. During the discussion, one of the students adopted a dominant role and 'over-participated.' The counsellor dealt with this most effectively and unobtrusively simply by interrupting the student once 'in order to hear what Mrs. X was trying to say'. Without any suggestion of impoliteness, without denying the value of the contribution of the dominant student, the point was made. The student, I thought, paused and considered rather more during the remainder of the discussion.

Resentment against the dominant student wasn't allowed to build up. He was encouraged to think a little more about his own behaviour. Other members of the group were encouraged not only directly, but also by perceiving the fundamental fairness of the counsellor, and the attention he was giving to the functioning of the group."

"He runs a loosely structured group. At first it's difficult to tell which is the counsellor. The discussion ambles along and occasionally strays. When it does, he lets it go to see if anything profitable is coming. If not, he brings it back in a way which causes no-one resentment. It is very hard to cut off a bore in midstream without diminishing him but he is able to do this.

For instance, on one occasion one man was generalising from the particular. The counsellor found something to praise in what the student had said, but ended up with a question which pointed him back to the subject. Other counsellors would say, 'You're straying off the

point, Tom. That won't do. Let's get back to the subject.' The group would probably support this, but Tom would probably be hurt badly and would find it difficult to come in again without fear of being put down. This counsellor is good at not putting people down."

In the case of ineffective counsellors, they sometimes either failed to develop group sessions or were unable to manage them in a profitable manner. Here are some examples:

"When I arrived at the study centre, there were six or seven students present. The group sat in a room with this counsellor and another, and they were going to discuss Hamlet. He was ponderous, he slapped students down rather aggressively, saying, 'Don't be childish' or telling a man of his own age 'not to be a silly boy'. He meant it as a joke, but it was nasty when the student was trying to express himself. He didn't lead the discussion. He let it wander from the subject and destroyed the atmosphere instead of building on it. I felt that the students would need to be resilient to return to the study centre. He certainly put me off."

"He knew he shouldn't lecture to his group. He knew he should lead a discussion. So he would start by asking a question which no-one could answer. He wasn't provocative. The students all sat in desks (it was a school) and they all went to the back of the room. I suggested that we might sit in a ring but he said that students in this group didn't need a ring. He said, 'I know it's the latest in group work

but these students are adults and don't need it.' I let it go but in fact the students were hiding behind their desks. It makes them feel secure but it's a barrier."

"This counsellor works at a small, rural study centre. He has a chemistry laboratory for a room. There were loose tables and chairs but he sat at the front on a podium. He could have easily reorganised the furniture but he didn't. I said 'Shall we change the shape of the room?' and he said 'No', though he couldn't give any reason. Later I said it would have made the room more amenable, but he couldn't see why it wasn't amenable as it was. There were eight people and forty seats.

Eventually, a fitful discussion began. They were discussing the A100 assignment and I have never sat through a more boring session in all my life. The assignment was concerned with poetry, and one working-class student had got to the heart of the poem but couldn't express himself fluently. When a middle-class woman shouted him down with fluent but superficial criticisms, the counsellor agreed with her, and said 'I think Mrs. X is right.' Who can be right about a poem? Every avenue which was presented which was away from right or wrong was dismissed by the counsellor. If a challenging idea was presented he would say 'Well I don't think we've got time for that.' Ideas were flowing but he didn't use them. He simply ruled them out."

"It was a bad night weather-wise and he was a bit embarrassed by the absence of students. He did nothing to make anything of the evening. Two or three students were present, chatting with the counsellor. A few more trickled in and it just went on as a kind of undirected conversation. It was linked to University activities, but there was no directed discussion. He made no move to get things going, no attempt to do anything constructive."

"He had the bad luck to have in his group one of the worst examples of the study centre bore, the student who haunts the study centre aggressively dominating every tutorial and self-help group. This man needed to be skilfully tamed by the counsellor if he was not to wreck the other students' work. He wasn't, and a whole year's work for some students was marred."

"I arranged the induction meeting so that students could go and meet their counsellors after the introductory talk. I told the counsellors to site themselves in various parts of the room so that I could direct students straight to them. All the students had to do was collect a cup of coffee and go and join their counsellor, who would give them a rundown and tell them about preliminary work and so on.

Good counsellors managed, with varying sized groups, to establish good relationships so that the meetings got going. But this counsellor's group was sitting in a rather chilly atmosphere. He was desperately trying to get something moving but just wasn't succeeding.

Later on I saw him at the study centre. There were two A100 groups with counsellors, and this counsellor. He was preparing to show a film but no-one came. He had already alienated the group. They didn't think it was worth coming any more."

#### Category 11

Effective counsellors were sometimes instanced as being academically capable whereas ineffective counsellors sometimes appeared to be academically weak. Moreover, effective counsellors were sometimes recommended for their academic approach, which demonstrated a recognition that OU teaching was not necessarily the same as teaching in other contexts. Sometimes, ineffective counsellors did not seem to make this distinction.

#### Effective counsellor examples:

"I visited the group when they were doing the section of A100 on history sources. The counsellor had mastered the material but he wasn't hitting it directly. He was trying to draw out from students their appreciation of source materials and their variety.

To underline this he brought in some diaries of an early 19th century country parson. The students could examine this material with some awareness of the main themes of 19th century history, and it gave them some insight into the different types of source material. The counsellor used this to pick up some themes in A100, the contrast of rural and industrial life.



He had set up an effective teaching situation involving students in the discussion of key concepts in the course at that point. Instead of focussing narrowly on the course he had brought in 'real' source material. He was operating as an academic counsellor - though I prefer to call it tuition."

"She has academic credibility. She can lead a discussion group and students will take tips from her on how to study. She can convey this to them without making them feel they can't attain such heights."

"I've seen him teach and the look of enthusiasm and excitement on the students' faces. He has great academic ability and flair. He excites the students in the subject, brings it out and makes it come alive. He knows the material and has a knack of focussing on interesting and exciting issues in the course."

"He has great mathematical expertise. His students quickly formed themselves into a group which has persisted far beyond M100 bounds. Inevitably the distinction between counselling and tutoring has become blurred."

"He reads the course material. He does not waffle. He is an excellent teacher and his students hold him in high regard, going to him for advice on matters both academic and administrative."

Ineffective counsellors examples:

"He doesn't prepare and he doesn't know the course - even though he gave an off-the-cuff talk on Goethe and Wordsworth which the other counsellors couldn't have done. He doesn't concern himself with the learning problems of students. I saw him discussing Hamlet once and he didn't seem to know the play."

"In 1972 she carried on as before but her abilities didn't show up so well. She hadn't got the academic background nor the insight into the interpersonal side of teaching. This began to make sessions a bit threadbare. As a colleague said, 'There are some counsellors who are worth their weight in gold in the first two weeks - their personality, understanding and ability to sort out problems give confidence to students. But they descend steeply once they get into the course material.' I don't think she had the ability to identify questions and help others to do so. She would always be helpful in telling people where to get help, but it never got into the sessions themselves."

"The Staff Tutor visited him and thought he was only rehashing the material he used in his daily teaching. He was delivering thirty-minuted patches which in his view were complementary to the course but which the Staff Tutor thought were irrelevant. He should have been trying to master the OU materials and base his approach on them."

"The first time I saw him at work I couldn't believe my eyes. There were six students present and they were discussing Hamlet. As I came in he jumped to his feet and introduced me in a mock-hostile way as 'The Powers That Be', and handed over to me. I told him that I

didn't want to interrupt and would just sit in as a member of the group. He said, 'Oh, do you want me to do something? Well, we're going to talk about Hamlet'. He took out a dog-eared set of notes and said, 'I've got this here. I haven't looked at it for a few years. It's what I did at college.' I couldn't believe it. He began to read from his despicable garbled notes. Every now and then he would stop and say, 'I suppose Shakespeare put this bit in because the play wasn't long enough'. It was the erection of philistinism into a system of education. He was ignorant and idle."

#### Category 12

Knowledge and understanding of the administrative regulations is mentioned in relation to effective and ineffective counselling. As we have already seen, understanding the administrative system influences the counsellor's ability to handle individual students' problems, and the kinds of difficulties a counsellor may refer to a Senior Counsellor.

The following remarks were made about effective counsellors:

"He knows the administrative cogs. He reads the material and the regulations."

"He is good on the system. He has taken the trouble to really familiarise himself with it and understand it."

"His 'administrative' work is done with meticulous accuracy and he shows a thorough grasp of the University's regulations and literature."

"He understands the OU system well and is administratively most efficient."

And for ineffective counsellors:

"He doesn't know the rules. Most counsellors don't read everything and I don't either. But he doesn't keep up. Students have asked him questions and he can't answer. He isn't the reference point for information which a counsellor should be."

"He felt our regulations weren't designed to meet with his inefficiency. He had really never read them. He had no command of the regulations."

"He never looks at the B.A. Degree Handbook."

### Category 13

Liaison with other part-time colleagues, especially tutors, was another activity mentioned with regard to effective and ineffective counsellors.

Examples of effective counsellors:

"They discussed things they were going to bring up at the tutorial. I thought she had previously been in touch with the course tutor. She outlined the things the tutor was intending to discuss. She didn't impose items but rekindled interest in those areas the tutor was going to cover. Her practice was to keep in close contact with tutors. She could have done other things such as refreshing past units or looking forward. Concentrating on what was to receive immediate specialist treatment was good."

"This counsellor works with another who is young but reticent and quiet. The counsellor organised his quiet colleague, who didn't resent it because it was a good idea and tactfully done. They arranged things so they would each take the other's Foundation-level students in their own area of academic competence. It was to the counsellor's credit that he disposed of the available forces in a productive way."

"He has seen the mistakes students made over conditional registration. He has a good working relationship with the Arts and Social Science counsellors, so they have been able to work together to ensure that students get a good overview of the courses."

"He has made good use of formal and informal meetings to establish good relations with full- and part-time OU staff."

Here are some examples illustrating a failure to liaise with part-time colleagues:

"I went to the study centre and the counsellors was there with some students. I asked him if his fellow-counsellor had arrived. He said he hadn't met him. I went next door and found the other counsellor. Neither had met the other. The counsellor knew his colleague's name and that he was on the same night but he hadn't bothered to meet him. He didn't even know that the study centre was supposed to be an entity."

"On this evening the counsellor had to leave early. He said he wasn't so unhappy to be going because the tutor was there to carry things through. He sat in for part of the tutorial and then left. He didn't make any notes about what was going on, or anything that might follow up. He had given no help at all to the tutor in the first place. I felt there was no link at all. I gather there was no contact afterwards with the tutor to find out what needed following up. There was no point in the counsellor being there, as far as I could see."

"I came along one evening to find three of the counsellors sitting in a room planning the month's programme. When I asked where this particular counsellor was, they said he was marking scripts in the tutorial room. Sure enough, he was sitting at the back, among the students, marking school-books. He got a bad mark from me. It was dereliction of duty."

#### Category 14

Failure to attend to the routine management of the study centre was occasionally mentioned with regard to eneffective counsellors. Good study centre management was rarely mentioned in respect of effective counsellors. This might suggest that if a counsellor did attend to study centre management he was not likely to be regarded as particularly effective. On the other hand, if he did not, this was likely to be regarded as a component of ineffective-ness. Nevertheless, for completeness this dimension seems worth including.

The few examples that follow therefore refer to ineffective counsellors:

"I arrived at the study centre at 6.30. I went into the building and I couldn't see any notices directing students. I went up in the lift to the third floor where our rooms are to check that the counsellor was in. He wasn't. There was nothing open and no-one there.

Although he was a member of the host institution, he didn't really worry about the state of the cupboards, adequate signposting or general facilities. He should have made sure that the rooms for tutorials were signposted as he had the information."

"There was some trouble concerning the keys of the study centre cupboard. He lost a set or didn't return a set, so other people couldn't get into the cupboard. This shows a lack of responsibility towards the place."

"There are no caretaking staff on duty during the evenings at the study centre where this counsellor operates. However, all counsellors have been issued with keys. This means that the counsellor on duty is the only person who is able to ensure that students have access to the OU cupboard, TV and computer terminal.

Some students travel fairly long distances in order to use the computer terminal, and I have had several reports that students have been unable to gain access to the terminal at 6.30 p.m. on evenings when this counsellor should have been present. The counsellor

arrived later to meet his own students, but this was little consolation to students who had 'booked' time on the terminal and who felt extremely frustrated when they couldn't get to it.

I have reminded him of his responsibilities in no uncertain terms, but I can never really feel confident that things will run smoothly on his evenings. He has been generally weak on study centre management; he has not always made sure that suitable accommodation is arranged for visiting tutors and, on occasion, he has gone away leaving lights on and doors unlocked."

## 5.0 Practical Methods of Selection and Supervision

A further by-product of the study concerns the methods of selection and supervision which are employed by Senior Counsellors in relation to counsellors. I have called these practical methods in that they represent the methods which some Senior Counsellors at least, utilise in practice. To some extent this may confirm what is already known about such matters.

### 5.1 Selection

The degrees of freedom which Senior Counsellors have over the selection of counsellors are circumscribed in various ways. Sometimes the choice is limited by logistical or academic factors which may mean that there is little choice. Some geographical and subject-matter areas may have a surplus of applicants while others do not.



In the selection process, the use of interviews seems to vary. Some Senior Counsellors said they always interviewed potential appointees whilst others were unable to do this. Instead they might speak to them on the phone, talk to their referees or make what inferences they could from the application forms. One Senior Counsellor made the following comments:

"It's hard to divine the characteristics of newcomers. Some people show an initial keenness which doesn't emerge in practice. I'm not sure interviews would show it. I think we almost have to have a trial year before we can tell whether these people will be any good. It means that when we look at the application forms we aren't looking for surface, obvious things but something which indicates character and attitudes. These are difficult things to pick out and may be displayed at any point on an application form - the carefulness and consideration with which it is filled in. If some parts aren't completed it tends to indicate that he doesn't consider the whole approach to the job very important. We have to try to pick up clues like that. I also try to get information through the grapevine."

## 5.2 Supervision

Once appointed the question of supervision arises. Senior Counsellors probably vary in the emphasis they place on this. One said:

"I see my job as to perform a service for my 38 counsellors, and to do the things they can't do. We accept that we are colleagues in a University. Part-time staff are full members of the University and I don't see the relationship as one in which I can order them about.

I don't see it as part of my role to be a kind of local authority inspector snooper, though it is part of the job and I do watch what counsellors do when I visit."

Another reported as follows:

"The amount of feedback you can get on counsellors' work partly depends on your own priorities. I attach considerable importance to handling problems referred to me by counsellors in detail. This takes time and a large slice of my working week. I've devoted more time to conscientious counsellors and students' problems than to deficient counsellors. It is important to act as you expect counsellors to act. You have to give the same effort and concern to problems as you expect the counsellor to give. My way of working doesn't enable me to monitor counsellors as closely as I would like."

Nevertheless, Senior Counsellors clearly do have to make judgements about the behaviour of their counsellors, and they are faced with some difficulties. These involve the amount of information Senior Counsellors are able to obtain on the work of their counsellors.

Opinions seem to differ as to whether enough information can be acquired to allow a reasonable judgement. Some Senior Counsellors felt that they did get enough: others felt they didn't. In any event, they have to rely on a number of clues and signals from a variety of sources. These sources may be divided into direct observations and indirect feedback via forms and other means.

### 5.2.1 Observation

Probably the most common form of this is via the study centre visit. The opportunities for this are limited by the number of counsellors 'attached' to a Senior Counsellor and the geographical features of the Region. As one Senior Counsellor said:

"I visited him three times in all, which is once more than the average when you have forty-five counsellors."

Even when a visit is made, certain aspects of the counsellor's work may not be open to observation. Thus individual interviews may be regarded as too sensitive or personal to allow observation without disturbance.

Observations made at the study centre may also be difficult to interpret. Thus:

"I only saw him a couple of times. Once I stayed for forty-five minutes. This was the main time. On the other occasion he wasn't with the group. When you go to the study centre it's the luck of the draw what you find people doing. How can you judge if someone is up to it academically? You only see them a few times. They may be having an off night."

Making judgements on academic grounds may be particularly difficult when the counsellor's subject is unfamiliar to the Senior Counsellor. Thus:

"You're like a spare part in counselling groups if you don't have the barest lay competence in the discipline .... I'm trying not to think only of Arts or Social Science counsellors whose work is obviously most rewarding for me to watch. I understand it better. I have good Science and Technology counsellors and, I suppose, Maths, but I'm not competent to assess their work in many ways. I can judge their personal/welfare work and can get an opinion from the Staff Tutor."

Other opportunities for observation may occur at large-scale meetings (such as briefing conferences), and small-scale meetings (such as social meetings).

#### 5.2.2 Indirect Feedback

The most common form of this is probably the C2 report form and the C5 attendance card. We have already seen (category 8) some of the ways in which C2 forms may be used as a judgemental device. Attendance records may also figure prominently in the Senior Counsellor's monitoring process. A counsellor who gets a consistently low attendance may become regarded as suspect. C3 and C4 forms may also be taken into account.

The remaining sources of feedback consist of reports from students, other full and part-time staff. Staff tutors and Senior Counsellor colleagues are obvious sources of reference, part-time staff less so, except perhaps in the case of complaints. As far as students are concerned, there may be some

variation in the extent to which their explicit reactions to a counsellor influence the Senior Counsellor's views of him. Thus:

"I do have some complaints about counsellors but in my experience they are mostly ill-founded, at least in the way they are expressed. For instance, one student told a Regional meeting that students at his study-centre thought counselling was a wash-out. It turned out that he was talking about his impression of his own counsellor, whose work I admire. The student was used to being told what to do and when he realised the OU wasn't like this he got upset. I told him that the counsellor was trying to make students more self-sufficient, and after a while he was prepared to see it this way. I spoke to the counsellor about it and he said he knew about it and had already had it out with the student. But the student was still adamant that he wanted it on a plate.

Many student complaints are of this kind - 'I'm not getting what I want'. A significant minority of students are working something out in the counselling relationship - it's the first time an adult has been paid to listen to them. I find that student complaints are of little help in assessing a counsellor."

Probably it is unlikely that a student will complain about a counsellor anyway.

"In general students are less sure about what they can expect from the counsellor and are less ready to come forward with complaints. With tutors they think they know what they should get more clearly. Anyway they are more anxious about tuition because of the grades."

"You are unlikely to get complaints from students. They don't have the confidence or knowledge to complain about the person we have said is their link with the system. The only time that happens is at summer school. Having talked to other students, a student might complain to the summer school counsellor that his own counsellor wasn't doing much and could something be done about it."

It is from sources such as those indicated above that the Senior Counsellor forms his opinion. Sometimes the way this was done was referred to as a "merging of experiences" which seems to express the process rather well. In some cases Senior Counsellors adopted a policy of selective attention to their counsellors. In the case of some established counsellors it was reported that visits to see them were rarely made since they were considered to be confidently in command of the role.

One Senior Counsellor provided a particularly good illustration of a judgemental strategy:

"I like to see counsellors at least once informally, though I usually can't fit them all in in a year. I like to see their wives and husbands. If it's a new counsellor I visit them at the study centre at least three times and have them round once for a meal.

At the first meeting there are usually a lot of apologies - 'I know it's not what you said at the briefing meeting but I think .....'. That's important. You can argue with 'I think's' but you can't argue with prejudices. You get odd things such as students' criticisms or bouquets. I don't take too much notice of these.

If I think it's good I may not make another visit unless I haven't seen a certain Faculty group for some time.

If it's a poor session I think, 'Maybe it was an off night' so I go again. If it's no better I decide to carry out some tests. I'll phone the counsellor up about a student and he probably won't know about him without the papers in front of him. Next time I make a visit, I warn the counsellor I'm coming and say I'm doing something concerning the registers - 'Can you have all yours together so I can look at them?' They're messy. There may seem to be no grip of the situation. I might say, 'You haven't marked X'. Some say, 'I haven't time to mark X' but they still know about X. Others say 'Oh well I must mark him in'. It's very impressionistic. The register isn't important in itself but the whole thing builds up. I think to myself 'He's dodgy'. So I go again and again, at least six times before I finally make up my mind.

I look for various indications like whether he excites the students. I look at the C2s and the withdrawal rates. I look at students' letters but they tend to influence me slightly in the opposite direction to the students' reaction.

It's an awful way of deciding but what else can we do?"

Senior Counsellors may build up a mental file of information about a counsellor and sometimes an explicit grading process may take place when the time to consider reappointment arrives. I have no information on how widespread this practice is, but in at least one Region a three point rating scale is used jointly by the Staff Tutor and Senior Counsellor. This grades the counsellor according to his suitability for reappointment.

In the event of a decision not to reappoint, some Senior Counsellors expressed a certain amount of embarrassment or difficulty about having to communicate this to people who may have assumed an implicit right to reappointment, although there are, of course, formal procedures for doing this.

It is hoped that this brief account will provide some insights into the problems and practices of selection and supervision. Some general comments will be made in the next section.

## 6.0 General Comments and Recommendations

This study was designed and undertaken largely before the Review Group on Tuition and Counselling began work. The relevance of any suggestions or recommendations is therefore partly dependent on the kinds of changes which may take place in the counselling service in the future. However, if the counselling service remains broadly as it is for the next two years, the following suggestions may be taken up within that period.



These comments are prefaced by the assumption that there is a measure of broad agreement as to the desirability and relevance of the characteristics stated above. If this is the case, then we can move on to consider how this material might be used to improve or encourage better counselling practice.

As I worked through the interviews and reports given to me by Senior Counsellors, I was increasingly aware of how difficult it must be for someone who is totally new to the OU to take on the counselling role successfully. The demands which could be made on a counsellor seem quite high. He may organise his sessions, familiarise himself with a complex set of regulations and procedures, follow the courses, develop relationships with individual students and monitor their progress, communicate with perhaps a number of tutors, make reports to the Senior Counsellor, look after the study centre and so on - and probably do a full-time job and run his social and domestic life, also. The person who can fully master all these aspects of the counsellor's role is probably a rare one.

There is one aspect of the circumstances in which the counsellor must work which seems worthy of mention. The counsellor must work in a rather loosely structured environment. By this I mean that the counsellor cannot depend on seeing each of his students regularly; he cannot assume that students' problems will be brought to his notice; he may not be able or wish to set out his own plan of activity - he may have to continually up-date his plans in the light of circumstances. In other words there is an element of unpredictability about the situations counsellors must deal with; and an element of concealment in that some areas of concern to the counsellor may not be automatically visible.

Effective counsellors seemed to respond to this situation by being active and making things happen. They are, perhaps, on the look out for opportunities to help students, actively searching out problems and solutions. An ineffective counsellor, on the other hand, might simply wait for things to happen, working in a purely reactive manner. The circumstances seem to make either approach possible.

#### 6.1 Selection

As has been noted above, the practice of interviewing applicants does occur in some Regions. If interviews are being conducted, the list of characteristics could be used as a check-list of areas to be considered by the interviewer. He may consider whether the applicant is likely to, or willing to, exhibit the kinds of behaviour referred to in each category. He may use the list as a self-checking device which reminds him of the range of abilities which may be important.

#### 6.2 Briefing and Training

a) Counsellors probably already receive a great deal of written briefing material such as the Handbook, Course Unit on Counselling and B.A. Degree Handbook. As an adjunct to this rather densely packed material, it may be valuable to produce a set of Important Points for counsellors. The aim would be to draw out and emphasise a limited number of important dimensions of counselling activity. This may serve as a kind of self-checking device for counsellors.

It would represent a series of peaks standing out from the horizon of counselling.

To produce this material I would suggest that Senior Counsellors, individually or together, take each one of the selected categories and write some working notes indicating:

Why this activity is important in counselling.

Suggestions as to how to do it.

Illustrations of the possible problems involved.

As an example here is some 'imaginary' material for category 9.

#### Programming Study Centre Activities

Importance: experience so far indicates that the programming of Study Centre activities is often desirable. This may take several forms:

1. Year-long programmes of course-group contact:- involves setting aside particular sessions for work with different course groups.
2. Year-long programmes of course-group work:- involves specifying areas or Units of the course(s) for consideration in particular sessions.
3. Session programmes:- involves dividing each session into periods for group work, individual consultation (etc.).

1. is important for ensuring that each course group is able to meet you for a group session on a regular basis.
2. is important because it motivates students to attend, enables them to prepare and enables them to select the times when they most feel they need to attend.
3. is important because it saves students time to know when you are available for individual face-to-face consultation.

#### What to do

For 1. define course groups from your allocation list. Consult fellow counsellors about possibility of amalgamating course groups. Produce contact programme. Circulate to students as a suggestion.

For 2. examine course timetable and seek information on the most common 'sticking-points' in the course. Allocate course work to sessions and circulate to students.

For 3. allow half-an-hour at the beginning and end of each session for personal consultation. Notify students, reminding them that you are also available for telephone consultation at home at specified times.

#### Problems

For 1. you may not have enough students following the same courses or disciplines to be able to form a comprehensive set of groups. Contact your fellow counsellors or the Senior Counsellor to see if any amalgamation is possible.

For 2. flexibility is necessary as students' needs vary throughout the year. You may have to work with groups following disciplines which are unfamiliar to you. You may not have course materials for post-Foundation courses in your own discipline (unless you are tutoring them) but copies are available for reference at the Study Centre.

For 3. - no major problems.

b) in verbal briefings, Senior Counsellors may like to use the list of characteristics and examples (in a suitably tactful way!) to illustrate some of the things which may be involved in good counselling.

### 6.3 Supervision and Monitoring

As has been noted above, the opportunities for monitoring counsellors' activities by Senior Counsellors is limited. Monitoring is generally useful when the question of reappointment arises, but it also seems to have two particular uses.

1. To enable encouragement and reward to be given to effective counsellors.
2. To enable assistance to be given to counsellors who may be in difficulty.

Before either of these things could be done, the Senior Counsellor needs to be able to get information about the counsellor's work. The logistical problems involved in doing this for all counsellors are very great, and even the desirability of attempting more

systematic monitoring and feedback may be questioned. In the case of letting counsellors know when they are highly thought of there may be little problem. The other situation may present problems. For instance the Senior Counsellor may find it difficult to establish that the counsellor does have weaknesses. He may be wary of trying to do anything in case the counsellor withdraws, leaving a gap to be filled. He may now know what to do about it anyway. And, probably, he will simply be too busy to be able to devote much time to a counsellor problem (as opposed to a counselling problem).

One approach might be to concentrate attention on new counsellors each year, on the assumption that a counsellor is most likely to have problems in his first year. This may happen already. The Senior Counsellor could use the list of characteristics as a crude check-list to help him systematically appraise the counsellor and to provide a framework for his observations. Senior Counsellors may also find it helpful to record samples of counsellor behaviour which they regard as important and as effective or ineffective. This is obviously useful when reviewing a counsellor's work.

## 7.0 Overview

The report began by describing the background and structure of the study of counsellor characteristics. Some problems involved in the study were reviewed. The main results of the study were

given in the form of a list of qualitative dimensions related to counselling, and a series of illustrative examples for each category. Then some information was provided on practical methods of supervision and selection. Finally some suggestions and recommendations were made for using the results in the processes of selection, briefing and training, and monitoring and supervision of counsellors.

Last, but not least, I wish to acknowledge the co-operation and effort given by Senior Counsellors in providing much of the information upon which this report is based.

ALAN THOMAS  
Research Assistant IET  
January 1974

## APPENDIX I

The table below shows the number of Senior Counsellors who cited an example of counselling behaviour within each of the categories of the list (pages 6-7 above). The number of counsellors who displayed behaviour within each of the categories is also shown.

There are no clear rules for deciding whether or not to include a particular category. The number of Senior Counsellors citing examples within a category has some influence but the lack of a large number of mentions was not regarded as grounds for excluding a category. This was because the aim was to get a fairly comprehensive set of categories, and thus to provide a set of general dimensions, all of which are of possible relevance.

Category	No. Senior Counsellors		No. of Counsellors	
	Effective	Ineffective	Effective	Ineffective
1	15	8	23	10
2	21	9	29	12
3	14	5	20	7
4	12	5	13	6
5	13	5	16	5
6	10	5	12	5
7	13	10	19	10
8	14	14	17	17
9	10	3	13	4
10	17	13	24	16
11	10	9	12	9
12	8	8	10	9
13	12	3	15	4
14	1	2	1	3



## APPENDIX II

The work of a total of 75 counsellors was reported on by Senior Counsellors. The figures below give some idea of the range of 'background' characteristics of these people. The information was often drawn from memory by Senior Counsellors and is therefore of a fairly rough and ready kind.

1. Sex

Male	60
Female	15

2. Age

Under 26	1
26-35	23
36-45	29
46-55	3
56-65	6
Over 65	1
No information	12

3. Course/Discipline

A100	13
D100	17
M100	4
S100	11
T100	7
Post-F-level	4
No information	19